View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by The University of Utah: J. Willard Marriott Digital Library

USING SPECIES DISTRIBUTION MODELS TO QUANTIFY CLIMATE CHANGE IMPACTS ON THE ROSY-FINCH SUPERSPECIES: AN ALPINE OBLIGATE

by

Edward C. Conrad

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Geography

The University of Utah

August 2015

Copyright © Edward C. Conrad 2015

All Rights Reserved

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of Edward C. Conrad

has been approved by the following supervisory committee members:

Simon C. Brewer	, Chair	04/16/15 Date Approved
Russell E. Norvell	, Member	04/16/15 Date Approved
Philip E. Dennison	, Member	04/16/15 Date Approved

and by Andrea R. Brunelle , Chair of

the Department of Geography

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Anthropogenic climate change is forcing plants and animals to respond by shifting their distributions poleward or upward in elevation. An organism's ability to track climate change is constrained if its habitat cannot shift and is projected to decrease in geographic extent. Geographic distributions were predicted for the Rosy-Finch superspecies (*Leucosticte atrata*, *Leucosticte australis*, *Leucosticte tephrocotis*) for climate change scenarios using correlative species distribution models (SDM). Multiple sources of uncertainty were quantified including choice of SDM, validation statistic, emissions scenario and general circulation model (GCM). Species distribution models are traditionally validated using a threshold-independent statistic called the area under the curve (AUC) of the receiver operating characteristic. However, during k-fold cross-validation, this statistic becomes overinflated due to spatial sorting bias existing between the presence and background points. Therefore, a calibrated area under the curve (cAUC) that removes spatial sorting bias, was used to assess predictive accuracy. Boosted regression trees (BRT) and random forest regression trees (RF) had significantly higher predictive skill than did maximum entropy (MaxEnt) or generalized additive models (GAM; $p = 0.05$). Predictions in space and time were made using the tree-based algorithms and applied to two representative concentration pathway (RCP) emissions scenarios, RCP2.6 and RCP8.5, for 15 and 17 GCMs, respectively, for 2061–2080. Despite variability in predictions, there is agreement between the projections for many regions suggesting that Rosy-Finches are extremely vulnerable to climate change. The short-term management implications of this study are the need for an immediate assessment of Rosy-Finches by surveying suitable habitats predicted by the species distribution

models to enable population estimates to be made and field validation of species distribution model outputs. Such an approach would revisit historical breeding sites lacking coverage in the eBird and Global Biodiversity Information Facility (GBIF) datasets to determine if some populations have already been extirpated due to climate warming. Long-term management implications are to consider the use of large herbivores in alpine and tundra ecosystems in order to mitigate ecosystem response to climate warming and preserve foraging habitat for Rosy-Finches during the breeding season.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF FIGURES vii

LIST OF ABBREVIATIONS viii

ACKNOWLEDGEMENTS ix

INTRODUCTION 1

- Background 4
- Emissions Scenarios 5
- Species Distribution Models – A Correlative Framework 6

METHODS 9

- Presence Data 9
- Environmental Data – Current Climate 10
- Environmental Data – Future Climate 13
- Statistical Procedures 14
- Statistical Models 15
- Boosted Regression Trees 16
- Random Forest Regression Trees 17
- Generalized Additive Model 18
- Maximum Entropy 19

RESULTS 24

- Species Distribution Model Accuracy 24
- Current Rosy-Finch Distributions 26
- Future Rosy-Finch Distributions 26
- Variable Importance 27

DISCUSSION 39

Limitations of Species Distribution Models 41
Implications for Management 44

CONCLUSION 49

APPENDIX: FUTURE CLIMATE CONSENSUS GRIDS 52

REFERENCES 61

LIST OF FIGURES

1. Nearctic Rosy-Finches 8
2. Rosy-Finch occurrences from eBird and GBIF 21
3. Rosy-Finch current distribution maps 30
4. Rosy-Finch current presence-absence maps 31
5. Rosy-Finch current agreement map 32
6. Rosy-Finch GCM agreement maps for 2061–2080 33
7. Rosy-Finch GCM consensus maps in 2061–2080 34
8. Projected change: Area and percentage for 2061–2080 35
9. Projected change: Gains, losses, and no change for 2061–2080 36
10. GCM consensus temperature anomalies (°C) for 2061–2080 47
11. Herbivory mitigates ecosystem response to climate warming 48
12. Future climate consensus grids for RCP2.6 52
13. Future climate consensus grids for RCP8.5 57

LIST OF ABBREVIATIONS

AUC	Area under the curve
AOU	American Ornithologists' Union
AR5	Fifth Assessment Report
BRT	Boosted regression trees
cAUC	Calibrated area under the curve
CMIP5	Coupled Model Intercomparison Project Phase 5
GAM	Generalized additive model
GBIF	Global Biodiversity Information Facility
GCM	General circulation model
GIS	Geographic information systems
IPCC	International Panel on Climate Change
MaxEnt	Maximum entropy
PRISM	Parameter-Elevation Relationships on Independent Slopes Model
RF	Random forest regression trees
RCP	Representative concentration pathway
SDM	Species distribution model

ACKNOWLEDGEMENTS

Pursuing a M.S. in Geography at the University of Utah has been a fulfilling process allowing me to explore related interests such as ecology, the impacts of climate change, geographical information systems (GIS), and statistical learning.

There are many people who have helped me along the way, but I am particularly indebted to my advisor, Simon Brewer, whose patience along the way has been remarkable. My introduction to the world of free and open-source software including R, GRASS GIS, GDAL Tools, Generic Mapping Tools, Perl and Unix-based scripting, and LaTeX is due to him. Not coincidentally, this has brought both joy and headache as I have discovered how tricky it can be to take simple conceptual ideas, and implement them in a rigorous analytical fashion. However, the headaches were entirely worth it and I have enjoyed the journey. Special thanks also goes to the two other members on my committee: Phillip Dennison and Russell Norvell. Each has provided valuable feedback and helped challenge me in regards to my choice in SDM algorithm and approach to model validation. I would also like to thank Kaitlin Barklow for teaming with me to create a GIS model in 2012 for the Black Rosy-Finch for our "Methods in GIS" project. This exercise got me excited about GIS and taught me that it was possible to model the niche of a Rosy-Finch.

I would lastly like to thank my parents and Caroline Jordan. My folks instilled in me a passion for the outdoors, and mountains in particular, by planning family vacations at national parks. Caroline, you have been instrumental to my success; for with you, I have a sounding board to articulate ideas, thoughts, and practice explanations.

INTRODUCTION

Correlative SDMs are commonly used as an aid for conservation planning, resource management, and environmental research (Franklin, 2010). However, SDMs have been criticized for being overly simplistic and for making potentially misleading predictions of species' geographic distributions, especially when used to project distributions under future climate scenarios (Beale, Lennon, & Gimona, 2008; Davis, Jenkinson, Lawton, Shorrocks, & Wood, 1998). The criticisms are partly warranted, because SDMs neglect to include data pertaining to an organism's morphology, behavior, and physiological tolerances to fully explain distribution limits, and therefore do not model a species' niche mechanistically (Buckley, 2008; Kearney, 2006; Kearney & Porter, 2009). Yet process-based models that are built upon ecophysiological relationships between an organism and abiotic variables are not a panacea to modeling species' distributions, because they too do not alleviate all uncertainties pertaining to factors that control species' distribution limits. Variability in predictions made by these two modeling frameworks is due to different assumptions made at the onset. Species distribution models model the realized niche by creating statistical relationships between observational data and the environment, which indirectly incorporates biotic interactions such as competition, since such processes contribute to geographic patterns in the observations dataset. Conversely, process-based models model the fundamental niche using ecophysiological relationships obtained by controlled experiments and therefore fail to account for biotic interactions at all. Since the realized niche is a subset of the fundamental niche, a consequence of making predictions in space and time is that SDMs tend to predict greater extinction risk than do process-based models (Morin

& Thuiller, 2009). Although SDMs are unable to account for phenotypic plasticity and local adaptation, their use as a tool in ecological research and conservation planning is justified since they do provide a means to predict species' responses to anthropogenic climate change and because few organisms are sufficiently well-studied to provide the requisite data to create process-based models.

To address the criticisms made regarding correlative SDMs and account for many of the uncertainties inherent in making predictions, a robust modeling framework is necessary. When mapping current geographic distributions, uncertainties exist at each stage of the modeling process from the choice of data inputs, to which SDM algorithm is used, to which resampling technique is used for model training and which validation statistics are used to measure predictive skill. When mapping future geographic distributions, additional uncertainties are introduced ranging from which emissions scenario is used, the time period being predicted, the choice of GCM, and how the GCM is statistically downscaled to a resolution useful for an ecological application. Therefore, in order for SDMs to be used to assess a species' risk to climate change, a robust modeling framework that quantifies many of these uncertainties is necessary to increase confidence in the predictions being made.

In this study, SDMs were constructed for the Nearctic Rosy-Finch superspecies, which comprises three closely related species: Black Rosy-Finch (*Leucosticte atrata*), Brown-capped Rosy-Finch (*Leucosticte australis*), and Gray-crowned Rosy-Finch (*Leucosticte tephrocotis*; Figure 1). These species are separated by geography, but breed exclusively on cliffs or steep talus slopes above treeline in the alpine-tundra biome (Johnson, 2002; Johnson, Hendricks, Pattie, & Hunter, 2000; Macdougall-Shackleton, Johnson, & Hahn, 2000).

There has been, and continues to be, considerable debate regarding the systematics of Nearctic Rosy-Finches due to the existence of eight allopatric forms

across North America. They have been treated as a single species as well as up to four species by various authors, but today are officially treated as three species — two monotypic species, *Leucosticte atrata* and *Leucosticte australis*; and a polygenic species, *Leucosticte tephrocotis* with six subspecies (Chesser et al., 2013). This taxonomic treatment is largely due to Johnson's (1973) unpublished dissertation titled, "Biosystematics of avian genus *Leucosticte*," in which he made extensive examination of study skins ($n = 2901$), took mensural measurements, and made extensive field observations, notably in a zone of overlap between *Leucosticte atrata* and *Leucosticte australis* where hybridization occurs in the Bitterroot, Cabinet, Little Belt, and Seven Devils Mountains in Idaho and Montana. Johnson (1973) claimed that sufficient breeding isolating mechanisms existed to prevent gene flow beyond this zone, and that plumage and morphological differences warranted separate species status for *Leucosticte atrata* and *Leucosticte australis* (Johnson, 2002). The American Ornithologists' Union (AOU, 1993) which officially holds responsibility for determining systematics of birds in North America, split the Rosy Finch, in-part upon these findings, and maintains this treatment despite genetic support, and reanalysis of Johnson's (1973) plumage and mensural data by Drovetski, Zink, and Mode (2009). The AOU North American Classification Committee deemed that further field and molecular research was necessary to support the conclusion made by Drovetski et al. (2009) to lump these three species (Chesser et al., 2013).

The avian genus *Leucosticte* occurs across the Holarctic, originates in the Palearctic, and includes four species restricted to Asia (Macdougall-Shackleton et al., 2000). Sibley and Monroe (1990) regarded the Nearctic Rosy-Finches along with the Asian Rosy-Finch (*Leucosticteg arctoa*) to be a superspecies, but molecular data from mitochondrial DNA restriction fragments ($n = 1$) supported a divergence event when birds of the genus *Leucosticte* first entered North America across the Beringia

land bridge (Zink, Rohwer, Andreev, & Dittmann, 1995). Further molecular research with larger sample sizes ($n = 201$) has since corroborated divergence of Nearctic Rosy-Finches from Palearctic Rosy-Finches. Although Chesser et al. (2013) did not lump the three Nearctic Rosy-Finches together as supported by the research done by Drovetski et al. (2009), it was not due to the committee's total dismissal of the findings. The reluctance shown by Chesser et al. (2013) to make any taxonomic changes was due to their recognition that there are close phylogenetic relationships between the birds and a prediction that further molecular research could reverse such a decision. Nearctic Rosy-Finches are classified as three species, but high uncertainty still remains regarding their systematics and supports my decision to model Nearctic Rosy-Finches as a superspecies (hereafter, Rosy-Finches). Such an approach also allows their current and future habitats to be quantified without subjectively defining boundaries necessary to train and test three species-specific SDMs. Furthermore, taxonomic controversies aside, Rosy-Finches share a vulnerability to anthropogenic climate change because they collectively breed at a higher altitude than any other species in North America (Johnson, 2002; Macdougall-Shackleton et al., 2000). A formal assessment of their current and predicted geographic distributions is therefore warranted, and makes them a prime candidate for demonstrating a robust SDM framework that may be applied towards controversial correlative SDM climate change applications.

Background

Plants and animals respond to a warming climate by shifting their distributions upslope in elevation or poleward in latitude. Observational data support these geographic shifts for many species of plants and animals. The data show that, although species do respond individualistically in the magnitude of change, they are already responding this way under postindustrial-era climate change (Parmesan, 1996; Parmesan et al., 1999; Parmesan & Yohe, 2003; Root et

al., 2003). Furthermore, pollen proxy data obtained from paleoecological research show that plants' primary responses to changing climates during the Quaternary Period have been to shift their distributions, rather than evolve and adapt to the change (Huntley, 1991; Huntley & Webb III, 1989; Prentice, Bartlein, & Webb III, 1991; Webb III & Bartlein, 1992). Predictive models are therefore required to examine future shifts in distributions due to climate change.

Understanding how species will respond to climate change is important because mean global surface temperatures have risen 0.78°C (0.72–0.85) from the 1850–1900 average to the 2003–2012 average (Stocker et al., 2013). In that period, CO_2 concentrations have risen 43% from 280 parts per million by volume (ppmv) in preindustrial times to 400 ppmv in 2013 — levels which are higher than any point in the past 800,000 years (Lüthi et al., 2008). Since there is a lag between a temperature increase and increased concentrations of greenhouse gases, climate scientists have intensely researched the Earth's equilibrium climate sensitivity in order to make accurate predictions of future climate warming. Presuming CO_2 emissions continue unabated, and the concentration of atmospheric CO_2 doubles compared with preindustrial levels, the most recent modeling research by the Coupled Model Intercomparison Project Phase 5 (CMIP5), concludes that there is "high confidence that equilibrium climate sensitivity is likely in the range 1.5°C to 4.5°C" (Stocker et al., 2013). Such an increase in mean global temperature will have significant ramifications for the Earth's biodiversity. For each 1°C rise in temperature, ecosystem zones are expected to shift 160 km poleward, or shift 160 m higher in altitude (Thuiller, 2007). SDMs provide scientists with a tool to forecast the individualistic responses of species to climate change (Thuiller, 2007).

Emissions Scenarios

Representative concentration pathways are plausible emissions scenarios developed by the global change community to address uncertainties in human

contributions to climate change and how Earth's climate will respond to future human and natural radiative forcings (Moss et al., 2010). As of 2011, total anthropogenic radiative forcing since 1750 has been 2.29 Wm^{-2} (Stocker et al., 2013). Four RCPs were developed based on different long-term socio-economic, environmental, and technological trends, which assume that policy actions will achieve one of the emission targets (Moss et al., 2010; Taylor, Stouffer, & Meehl, 2012). Representative concentration pathways provide time-varying concentrations of greenhouse gases for climate change research and impacts assessments and were adapted by CMIP5 (Taylor et al., 2012). The change in radiative forcing of each RCP is relative to preindustrial conditions, and is named for the radiative forcing expected by 2100. RCP8.5 is the high emissions scenario assuming an increase of 8.5 Wm^{-2} due to >1,370 CO_2-equivalent global warming potential; RCP4.5 and RCP6 are two intermediate scenarios with 650 and 850 CO_2-equivalent, respectively; and RCP2.6 is a scenario that peaks at 490 CO_2-equivalent prior to 2100 and then declines (Moss et al., 2010). Although the RCPs are to be treated as equally plausible outcomes, not predictions, RCP2.6 falls below the 10th percentile of mitigation scenarios while RCP8.5 represents the 90th percentile of emission scenarios in the published literature as of September 2007 (Moss et al., 2010). Therefore, the use of these two RCPs helps capture uncertainty in emissions scenarios and delimits a probable interval of future radiative forcing due to the combustion of fossil fuels.

Species Distribution Models — A Correlative Framework

Species distribution models, also known as ecological niche models, habitat suitability models, and climate envelope models, use georeferenced point data, representing an organism's presence, presence-absence, or abundance, and make correlations with environmental variables related to climate, topography, soil type, etc., occurring at those points. The model's output is a probabilistic map where

values indicate habitat suitability for the modeled organism. Species distribution models have become popular due to minimal data requirements together with improved data availability, and the advancement of GIS, which has facilitated easier handling of geospatial datasets, allowing detailed distributions maps to be made for any organism (Franklin, 2010; Guisan & Zimmermann, 2000).

The correlative approach to mapping species' geographic distributions is supported by ecological niche theory. The ecological niche was originally defined by Grinnell (1917) to be the place where environmental variables (e.g., temperature and precipitation), as well as habitat (i.e., land cover) could support a particular species. This is to say that Grinnell's definition of the niche referred to the environmental requirements. In contrast, Elton (1927) proposed an alternative definition of the ecological niche placing emphasis on the functional role a species has within its community — its food and interactions with other species (i.e., its impacts on the environment). However, it was Hutchinson's (1957) concept of the niche that now provides the theoretical framework for correlative SDMs. Hutchinson coupled Grinnell's and Elton's concepts, transforming the niche to apply to aspects of a species rather than to aspects of an environment (i.e., each species has its own niche rather than a habitat type has niches to be filled; Colwell & Rangel, 2009). In his "Concluding Remarks," Hutchinson (1957) introduced the fundamental niche as the "n-dimensional hypervolume," representing all variables — both abiotic and biotic — allowing a species to survive and reproduce in the absence of other species. He distinguished this from the realized niche, which he described as the subset of the fundamental niche occupied by a species due to interactions with competitors and predators. "Hutchinson's duality," which allows the niche to be described in environmental space and mapped in geographic space, has provided SDM modelers with a powerful means to conceptualize and analyze species' distributions in relation to environmental variables (Colwell & Rangel, 2009).

(a) (b) (c) (d) (e)

Figure 1. Nearctic Rosy-Finches. Five of the eight allopatric forms are depicted above. (a) Black Rosy Finch (*Leucosticte atrata*); (b) Brown-capped Rosy-Finch (*Leucosticte australis*); (c) Gray-crowned Rosy-Finch "Coastal" subspecies (*Leucosticte tephrocotis littoralis*); (d) Gray-crowned Rosy-Finch "Pribilofs" subspecies (*Leucosticte tephrocotis umbrina*); (e) Gray-crowned Rosy-Finch "Interior" subspecies (*Leucosticte tephrocotis tephrocotis*). Photos by Jacob Spendelow. Used with permission.

METHODS

Presence Data

Georeferenced occurrence records were acquired from GBIF, a free and open source biodiversity database (GBIF, 2012). However, the majority of Rosy-Finch occurrences acquired through GBIF, originated from eBird, a National Science Foundation funded avian database, jointly implemented by the Cornell Lab of Ornithology and the National Audubon Society (Sullivan et al., 2009). Therefore, in order to obtain the most-recent edition of the database, I contacted eBird directly and occurrence records through August 2014 were used (eBird, 2014). Such data represent opportunistic and/or purposive observations rather than those coming from probabilistic sample designs, yet the data are freely available, comprise the largest avian database in existence, and are available at a fine-grain spatial scale (i.e., latitude-longitude pairs).

Prior to use in SDMs, the occurrences were individually reviewed to improve model performance. This is an underappreciated facet of SDMs, but the quality of input presence-absence data has been shown to be a more important factor determining predictive skill than predictive power of the environmental grids or the SDM chosen (Lobo, 2008). Most importantly, I made a subselection of the data to include only records within the approximate breeding season: June–September. Rosy-Finch breeding phenology does in fact start earlier in the Bering Sea island populations, with birds laying eggs 28 April–26 May (Macdougall-Shakleton et al., 2000); however, records from the Aleutians and Pribilofs represent populations that inhabit Arctic tundra year-round. By restricting my records between the months of June through September, I was able to focus this analysis on the breeding and

foraging stage of Rosy-Finch life-history when all populations inhabit the alpine-tundra biome, at which time, they are the most vulnerable to climate change. In fact, there are zero breeding records away from this biome (R. Johnson, personal communication, February 25, 2012). Further edits were made to the occurrence dataset so that no duplicates existed for a single georeferenced occurrence point, nor did multiple unique geographic coordinates occurring within the same 828 m^2 grid cell (environmental grids discussed below). In some instances, particularly in the Aleutians and Pribilofs where Arctic tundra and suitable nest substrate occurs at sea level and irregular coastlines are common, georeferenced points had to be moved slightly from overlapping ocean to the nearest grid cell overlapping land due to the coarseness of the grids. Lastly, metadata for the type and locality of the records was reviewed, to check and fix georeferencing errors, as well as to remove vague records lacking geographic precision (e.g., state county centroid points from submitted eBird checklists or eBird traveling count checklists >10 km in length). After these steps, there were 1,395 unique Rosy-Finch occurrence records available for modeling (Figure 2).

Environmental Data — Current Climate

Twenty-one environmental grids were acquired from three sources and processed using GDAL 1.1.1 and GRASS GIS 6.4.4 (GDAL Development Team, 2014; GRASS Development Team, 2012). The grids were projected to the Albers Equal-Area projection, and clipped to the study extent, which encapsulated the entire distribution of the Rosy-Finch superspecies. Grids comprising continuous values were processed using bilinear resampling while categorical grids were processed using nearest-neighbor resampling. This resulted in 21 grids with a resolution of 828 m^2.

Climatic data were downloaded from WorldClim at 30 arc-second resolution and consisted of 19 bioclimatic variables (Hijmans, Cameron, Parra, Jones, &

Jarvis, 2005). Variables included annual mean temperature, mean diurnal temperature range, isothermality, temperature seasonality, maximum temperature of warmest month, minimum temperature of coldest month, temperature annual range, mean temperature of wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, mean temperature of coldest quarter, annual precipitation, precipitation of wettest month, precipitation of driest month, precipitation seasonality, precipitation of wettest quarter, precipitation of driest quarter, precipitation of warmest quarter, and precipitation of coldest quarter. WorldClim grids use the 1950–2000 climatic average and are created using the ANUSPLIN software package by an interpolation method involving thin-plate smoothing splines using latitude, longitude, and elevation as predictors (Hijmans et al., 2005). Known deficiencies regarding WorldClim include that it underestimates precipitation by as much as 25% in regions with complex topography, such as the western U.S. (Daly et al., 2008). Furthermore, the spatial distribution of the climate stations WorldClim uses is biased toward low elevations in higher latitudes, particularly the northwest United States and Canada, and biased towards higher elevations towards the tropics (Hijmans et al., 2005). Parameter-elevation relationships on independent slopes model (PRISM) climatic data better account for physiographic features in the interpolation process than does WorldClim, thereby providing higher accuracy in regions with much topography such as the western United States. However, WorldClim was chosen since the study extent of this analysis exceeds the spatial extent of the PRISM dataset (Daly et al., 2008). Lastly, for those grids that had decimals removed to reduce online data storage costs, a processing step was performed to restore these using map algebra.

A single biome grid at 30 arc-second resolution was obtained from Rehfeldt, Crookston, Sáenz-Romero, and Campbell (2012). This dataset consists of 46 modeled biotic communities created using random forest classification trees, based

upon the digitized biotic communities of Brown (1998) as input. Overall classification error averaged 3.7% in this model and for each biotic community, classification error was attained from 10 samples of 2000 observations (Rehfeldt et al., 2012). However, 76.4% of the total error occurred due to misclassifications of biotic communities not proximal to alpine or tundra, meaning less than 1% of overall classification error occurred for the biome of interest to the present study (Rehfeldt et al., 2012). For the purposes of the present study, 46 biotic communities were reclassified to a binary grid to indicate suitable and unsuitable biome with all alpine and tundra biotic communities classified as suitable and other biotic communities classified as unsuitable per Rosy-Finch breeding ecology. Of special note, the extent of the grid provided by Rehfeldt et al. (2012) did not extend beyond the International Date Line. Therefore, the biome grid was extended and classified as suitable to include the Commander Islands, so that the westernmost range of the Rosy-Finch according to the AOU could be included in the analysis (i.e., an assumption was made that the algorithm would have classified this region as Arctic tundra, as is found at the neighboring Aleutians, had it been included).

The last environmental grid was cliffs so that potential nesting sites could be included in the SDMs. Digital elevation models at 7.5 arc-second horizontal resolution (i.e., 225 m resolution) were downloaded from the Global Multi-Resolution Terrain Elevation Data (United States Geological Survey, 2010). The USGS (2010) data product used was spatially aggregated using the "breakline" emphasis and has a vertical accuracy as measured by root mean square error of 29 m. Slope was calculated in degrees using a 3x3 cell neighborhood around each input cell except for cells adjacent to the edge of the grid, which were assigned a slope of zero (GRASS GIS defaults). A threshold of 30° or greater was chosen to represent cliffs because this was the threshold necessary to capture suitable cliffs on this digital elevation model. A buffer of 2 km was created from the center of a cliff cell

to the centers of surrounding cells using Euclidean distance to represent proximal foraging area per Rosy-Finch breeding ecology (Johnson, 2002), as well as take into account georeferencing uncertainty in the original occurrence records. Therefore, cells meeting the cliff criteria or that were within 2 km of a cliff were reclassified as suitable and cells not meeting either of these criteria were reclassified as unsuitable. Finally, the grid was resampled to the same resolution as the coarser grids using the "r.resamp.stats" function, weighted for area, and using the maximum option (GRASS Development Team, 2012). Therefore, a cell classified as suitable at 7.5 arc-second would also be classified as suitable when resampled to 30 arc-second resolution or approximately 828 m^2 resolution when projected.

Environmental Data — Future Climate

Future climate grids were downloaded for two emissions scenarios: RCP2.6 and RCP8.5 for a single time period, 2061–2080. WorldClim statistically downscaled the grids, by taking the absolute (for temperature) or relative (for precipitation) difference between each GCM run for the 1961–1990 reference period and the GCM's run for the 2061–2080 target period and resampling the temperature or precipitation anomaly to 30 arc-second resolution. Next, WorldClim calibrated the output by taking the anomaly and applying it to their interpolated temperature and precipitation grids for current climate. WorldClim then derived the 19 bioclimatic variables for each GCM. For this study, 608 future climate grids were processed using GRASS GIS and GDAL scripts so that the projection, origin, extent, and spatial resolution matched that of the current climate grids. Grids for suitable biome and cliffs were the same as for current climate.

Species distribution models created using different GCMs produce different projections. Therefore, in order to capture GCM variability and be able to create projection agreement and consensus maps, all GCMs from the International Panel on Climate Change's (IPCC) Fifth Assessment Report (AR5) for the scenarios and

timeframe describe above that had been statistically downscaled and calibrated to WorldClim data, were used in this analysis. Fifteen GCMs were available for RCP2.6 while 17 GCMs were available for RCP8.5 for the 2061-2080 timeframe (Table 1).

Statistical Procedures

The modeling procedure for each SDM followed the same modeling framework and all statistical analyses were performed in R (R Core Team, 2014). Once the presence data and 21 environmental grids had been checked for errors and projected to the Albers Equal-Area projection, 10-fold cross-validation was used to create training and testing datasets using 1,395 presences and 10,000 randomly chosen background points that by chance could contain presence points. Since my objective for using SDMs was to have high predictive skill, model validation was done using a threshold-independent metric, the AUC of the receiver operating characteristic. This statistic is commonly used by SDMs, as it accounts for all possible information in the classifier, unlike threshold-dependent validation metrics (Fielding & Bell, 1997). Receiver operating characteristic plots depict all combinations of sensitivity (proportion of presences predicted correctly) along the y-axis and 1-specificity (proportion of background points incorrectly predicted as presences) along the x-axis (Fielding & Bell, 1997). The AUC ranges from 0.5 to 1 and an AUC of 0.95 indicates that 95% of the time, a dichotomous classifier will correctly order a randomly selected presence location and randomly selected absence location (Fielding & Bell, 1997; Phillips, Anderson, & Schapire, 2006). Therefore, an AUC value of 0.5 indicates that the classifier fails to order the positive case above the negative case and can be interpreted as a random prediction (Fielding & Bell, 1997; Phillips et al., 2006). For presence-only data models in which background points are used rather than true absence points, the AUC is still reported, but the interpretation changes to a dichotomous classifier's ability to order a randomly

selected presence above a randomly selected background point (Phillips et al., 2006).

A second validation step was performed to account for spatial sorting bias introduced to the AUC. This is to say that the geographic distance between training-presence and testing-presence points tends to be closer than the distance between training-presence and testing-background points, causing traditional AUC scores to become inflated (Hijmans, 2012). After creating a testing dataset with spatial sorting bias removed, model validation was again performed to calculate a cAUC. The cAUC was thereby regarded as the true validation procedure and used to rank and select SDMs for making climate change projections.

Statistical Models

It is important to note that species occurrence data from opportunistic or purposive observations represent a presence-only biological data model, due to imperfect detectability of vagile Rosy-Finches (MacKenzie et al., 2002; MacKenzie, Nichols, Sutton, Kawanishi, & Bailey, 2005). Although specialized modeling approaches to account for imperfect detection do exist (i.e., modeling occupancy rates), these approaches are used for large-scale monitoring programs or to investigate metapopulation dynamics (MacKenzie et al., 2002). Accounting for imperfect detection has been accomplished in SDMs using hierarchical occupancy models, with slight improvements in accuracy (Rota, Fletcher, Evans, & Hutto, 2011); however, since my objectives are to determine the response of Rosy-Finch's geographic distribution to climate change, and accounting for imperfect detection requires a dataset with survey points repeated over the course of a season, a presence-only SDM approach must be selected whereby statistical relationships between values of environmental grids at presences and background points are made.

Numerous statistical models and machine learning methods exist for constructing correlative SDMs. To address SDM uncertainty prior to making climate change projections, SDMs were created using four different algorithms:

BRT, GAM, MaxEnt, and RF. Although this is not an exhaustive list of SDM approaches, it does include three of the best SDMs at making predictions (BRT, MaxEnt, and RF) as determined by model comparative studies (e.g., Elith et al., 2006; Lawler, White, Neilson, & Blaustein, 2006), as well as a more classical statistical approach (GAM) useful for both explanatory and predictive purposes. Three of the SDMs fall in the supervised statistical learning domain because there is a response variable — presence or background — that is being predicted while GAM probability estimates are computed using maximum likelihood.

Boosted Regression Trees

A BRT model is a nonparametric machine learning algorithm that uses an ensemble of decision trees, and has higher predictive skill and lower variance compared to decision trees. In the context of SDMs, a BRT model works by fitting a single decision tree without pruning to predict a binomial response, presence or background (a binomial distribution is specified at the outset). However, a statistical procedure called boosting is then performed, in which subsequent trees are fitted using the residuals of the previous tree as the response. Thus, a BRT model is built sequentially, so that after each additional tree is fitted to the residuals by least-squares, the residuals are updated and the next tree is fitted to them (Friedman, 2002). The performance and computation speed of a BRT model is further improved by incorporating randomization into fitting each tree by randomly subselecting the training data without replacement (Friedman, 2002).

Several parameters control how a BRT model is built: the learning rate (shrinkage), tree complexity, bagging fraction, and number of trees. Generally, learning rate is set between 0.01 and 0.001 so the contribution of any single tree is small (James, Witten, Hastie, & Tibshirani, 2013; Ridgeway, 2013). Tree complexity determines the number of splits in each tree—for instance, a BRT model with a tree complexity equal to 1 consists of many trees with stumps. Tree complexity can also

be viewed as the interaction depth of the model, and its value, along with learning rate, determine how quickly the model learns to predict the response (James et al., 2013). The bag fraction is set to a value less than 1, although the optimal bag fraction value is dependent on a given dataset (Friedman, 2002). A bag fraction of 0.5 means that for any given tree, only half the training data are randomly selected without replacement. Lastly, it is possible to overfit a BRT model using too many trees, so the number of trees is determined using cross-validation (Ridgeway, 2013).

A BRT model was created with all predictors except for two climatic variables — temperature annual range and mean temperature of driest quarter — which had zero influence in all BRT models tried, and fitted with a tree-complexity of 10, a learning rate of 0.01, bag fraction of 0.6, and 3,450 trees that were determined by 10-fold cross-validation using the "dismo" package (Hijmans, Phillips, Leathwick, & Elith, 2014). Boosted regression tree models of various tree complexities, learning rates, and bag fractions were fitted, but the final chosen parameters were selected using the models with the highest cAUC scores.

Random Forest Regression Trees

A RF model is a nonparametric machine learning algorithm that builds many training datasets using the bootstrap, then fits a single regression decision tree for each without pruning, and takes the average across all trees. Creating a bootstrapped sample to train and test each tree, and doing this 100s–1000s of times, improves predictive skill, reduces variance, and is called bagging. A RF model, however, improves upon bagging by also limiting the number of predictors available that can be used for each split, which decorrelates the trees, thereby further reducing variance (James et al., 2013). For RF (as opposed to random forest classification trees), the number of randomly chosen predictors (m) available as candidates at each split is typically $m = p/3$, where p is the total number of predictors ($p = 21$ for this study), although this may manually be checked for a

particular dataset using the "tuneRF" function in the "randomForest" package (Liaw & Wiener, 2002). A RF model is not prone to over-fitting because the error rate converges asymptotically to a limit as the number of trees increases (Breiman, 2001). In practice, the number of trees chosen is that which is sufficiently large to allow the error rate to settle (James et al., 2013).

A RF model comprising 1000 trees using all 21 predictors (no predictors had zero influence) was created with seven randomly sampled predictors available at each split using the "randomForest" package (Liaw & Wiener, 2002). The "tuneRF" function confirmed that setting the "mtry" parameter to seven was optimal for reducing overall prediction error.

Generalized Additive Model

Logistic regression GAMs are semiparametric models that have more flexibility than logistic regression GLMs because they allow nonlinear relationships between continuous variables and the response using smoothed curves. However, like parametric models, an assumed probability distribution based on the type of response variable being modeled is provided. In the case of SDMs where the response has two possible values, presence or background, a binomial distribution is specified allowing a nonnormal error structure. Logistic regression GAMs use the logit as a link function that establishes a relationship between the smoothed function and the mean of the response (Guisan, Edwards Jr, & Hastie, 2002). Unlike the tree-based modeling approaches, however, GAMs cannot fit smoothed nonlinear relationships for categorical variables so they are included as dummy variables.

A GAM model was created using all 21 predictors with smoothing splines for the 19 continuous variables and the cliffs and biome variables treated as dummies. The logit function linked the predictors to the binomial response using the "mgcv" package (Wood, 2011).

Maximum Entropy

The machine learning algorithm, MaxEnt, makes no assumption of perfect detectability, predicts probability of occurrence, and is appropriate for presence-only datasets (Phillips et al., 2006). Entropy in statistics and information theory is a measure of dispersedness due to the amount of information in a probability density function. A probability density function containing higher entropy has less information (more uncertainty), so its probability distribution is more spread out and closer to uniform. The MaxEnt algorithm calculates a probability distribution for presence of an organism by creating two separate probability density functions: one for the presences $f_1(z)$ and one for the randomly sampled background $f(z)$, which by chance may contain presences (Elith et al., 2011). The probability density functions are created using six types of features: linear, quadratic, product (i.e., interactions among predictors), threshold (i.e., step-functions), category (i.e., categorical variables), and hinge (i.e., piecewise linear responses; Phillips & Dudík, 2008). Maximum entropy minimizes the relative entropy in covariate space and initially the ratio $f_1(z)/f(z)$ is calculated; however, this is next converted to a logistic output ($log f_1[z]/f[z]$) in order for probability of presence to be interpreted (Elith et al., 2011). Maximum entropy uses L_1 regularization to prevent overfitting, which keeps the complexity of the model from going beyond the empirical data (Phillips et al., 2006; Phillips & Dudík, 2008). Since the initial release of the Java-run MaxEnt software program, Phillips and Dudík (2008) have empirically tuned the regularization parameters by feature class with various presence-only sample sizes in a study using 266 species comprising four taxonomic groups from six regions of the world. Although in order to find the most optimal regularization parameters, these would need to be empirically tested on a species-by-species basis, Phillips and Dudík (2008) showed that these now established regularization defaults performed nearly as well for independent datasets. Lastly, similar to GAMs, but

unlike the tree-based models, MaxEnt is unable to create nonlinear response functions for categorical variables so these are treated as dummies.

A MaxEnt model using all feature types (i.e., linear, quadratic, product, threshold category, and hinge) was created with the default parameter values (iterations = 500, convergence = 0.00001, and regularization value = 1.0) using the "dismo" package (Hijmans et al., 2014). Maximum entropy models with less flexible response functions were fitted (hinge and categorical features only; linear, product, and categorical features only), but had lower predictive skill, as expected.

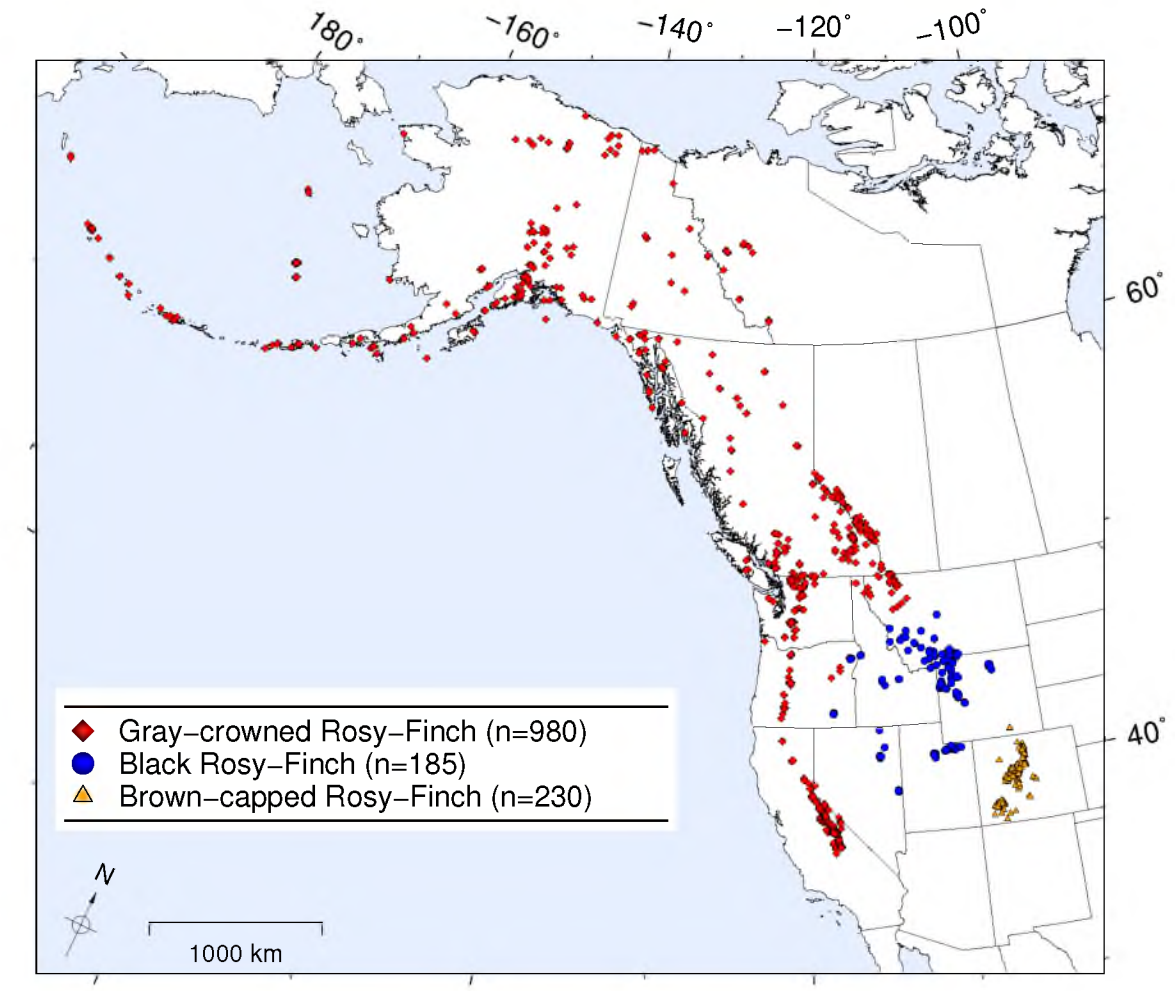


Figure 2. Rosy-Finch occurrences from eBird and GBIF. Occurrences (June-September; $n = 1,395$) used for training and testing SDMs by 10-fold cross-validation. Occurrences are from the three Nearctic Rosy-Finches (Black, Brown-capped, and Gray-crowned), but were combined for modeling the avian genus *Leucosticte* as a superspecies. GBIF = Global Biodiversity Information Facility; SDM = species distribution model.

Table 1

General circulation models from the Coupled Model Intercomparison Project Phase 5 used in this study

		RCP2.6		RCP8.5	
Modeling center	GCM	Warming[a]	Precipitation[b]	Warming[a]	Precipitation[b]
CSIRO-BOM	ACCESS1.0	–	–	5.75°	23.30
BCC	CSM1.1	2.34°	34.68	5.13°	63.23
RSMAS	CCSM4	2.33°	22.62	5.03°	43.59
CNRM-CERFACS	CM5	2.42°	38.96	4.69°	60.79
NOAA GFDL	CM3	4.01°	53.38	6.70°	96.50
NOAA GFDL	ESM2G	1.50°	32.69	–	–
NASA GISS	E2-R	1.29°	18.29	3.44°	32.61
NIMR/KMA	HadGEM2-AO	2.57°	27.83	6.52°	45.36
MOHC	HadGEM2-CC	–	–	6.65°	55.18
MOHC	HadGEM2-ES	3.28°	35.55	6.84°	47.07
INM	CM4	–	–	3.72°	48.83
IPSL	CM5A-LR	2.52°	27.85	6.03°	44.06
MIROC	ESM	3.43°	41.01	6.55°	66.65
MIROC	ESM-CHEM	3.73°	40.45	6.79°	64.69
MIROC	MIROC5	2.58°	25.04	5.27°	42.21
MPI-M	ESM-LR	2.37°	24.62	5.31°	52.38
MRI	CGCM3	1.33°	27.98	3.27°	61.89
NCC	NorESM1-M	2.40°	20.85	5.13°	36.56

Note. [a]Warming is the temperature anomaly (°C) for each GCM run from the 1961-1990 baseline to the 2061-2080 projection, statistically downscaled to 30 arc-second resolution and calibrated to the 1950–2000 WorldClim temperature girds (i.e., mean annual temperature). Next, the mean temperature over the entire geographic extent of my analysis was calculated in GRASS GIS using "r.stats" for each GCM, and the difference was calculated between it and the mean temperature of the current WorldClim mean annual temperature grid. This difference (i.e., the anomaly) is reported in the table. [b]Precipitation is the change in precipitation (mm/year) for each GCM run from the 1961-1990 baseline to the 2061-2080 projection, statistically downscaled to 30 arc-second resolution and calibrated to the 1950–2000 WorldClim total annual precipitation grid. Next, the mean value of total annual precipitation over the entire geographic extent of my analysis was calculated in GRASS GIS using "r.stats" for each GCM, and the difference was calculated between it and the mean total annual precipitation of the current WorldClim total annual temperature grid. This difference (i.e., the anomaly) is reported in the table. GIS = geographical information systems; GCM = general circulation model; RCP = representative concentration pathway; CSIRO = Commonwealth Scientific and Industrial Research Organization; BOM = Bureau of Meteorology (Australia); BCC = Beijing Climate

Center (China Meteorological Administration); RSMAS = Rosenstiel School of Marine and Atmospheric Science (University of Miami); CNRM-CERFACS = Centre National de Recherches Météorologiques Centre Européen de Recherche et Formation Avancée en Calcul Scientifique; NOAA = National Oceanic and Atmospheric Administration; GFDL = Geophysical Fluid Dynamics Laboratory; NASA = National Aeronautics and Space Administration; GISS = Goddard Institute for Space Studies; NIMR = National Institute of Meteorological Research; KMA = Korea Meteorological Administration; MOHC = Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais); INM = Institute for Numerical Mathematics; ISPL = Istitut Pierre-Simon Laplace; MIROC = Model for Interdisciplinary Research on Climate; MPI-M = Max-Planck-Institut für Meteorologies; MRI = Meteorological Research Institute; NCC = Norwegian Climate Centre; ACCESS = Australian Community Climate and Earth System Simulator; CSM1.1 = climate system model version 1.1; CCSM4 = community climate system model version 4; CM5 = climate model version 5; CM3 = coupled physical model version 3; ESM2G = Earth system model version 2 (generalized ocean layer dynamics); E2-R = model E distribution (coupled to the Russell ocean model); HadGEM2-AO = Hadley global environmental model 2 — atmosphere and ocean; HadGEM2-CC = Hadley global environmental model 2 — carbon cycle; HadGEM2-ES = Hadley global environmental model 2 — Earth system; CM4 = climate model 4; CM5A-LR = climate model 5 version A running on low resolution grid; ESM = Earth system model; ESM-CHEM = Earth system model coupled with atmospheric chemistry; MIROC5 = Model for Interdisciplinary Research on Climate 5; ESM-LR = Earth system model running on low resolution grid; CGCM3 = coupled global climate model 3 series; NorESM1-M = Norwegian Earth system model (core version of the model).

RESULTS

Species Distribution Model Accuracy

Accuracy of SDMs was measured using both threshold-dependent and threshold-independent validation metrics. Predictive skill was high across all four SDMs as measured by AUC, but the BRT model was statistically higher than the other three models ($p = 0.05$; Table 2). However, in order to assess a SDM, multiple validation criteria should be used (Elith & Leathwick, 2009). Confusion matrices, which cross-tabulate the observed and predicted values for a binary outcome, were created for each SDM using a threshold of 0.5 with the "PresenceAbsence" package (Freeman & Moisen, 2008b). Care must be taken when assessing a model's accuracy, since high accuracy may be attained simply if the prevalence of a species is low (Fielding & Bell, 1997; Franklin, 2010; McPherson, Jetz, & Rogers, 2004). For instance, the percentage of background points (i.e., true negatives) correctly predicted was high across all SDMs; however, for an organism with low prevalence, such as the Rosy-Finch, this is to be expected (Table 2). Since such model accuracy metrics are sensitive to the prevalence of a species, a more robust threshold-dependent metric should be used. Cohen's kappa (hereafter called kappa) is a threshold-dependent statistic that ranges from 0 to 1, uses all information in the confusion matrix (True Positives [TP], True Negatives [TN], False Positives [FP], and False Negatives [FN]), and is a metric that measures a classifier's improvement above chance (Equation 1; Fielding & Bell, 1997). Although it is inappropriate to compare kappa values across models with different species prevalence, because it too is affected by prevalence (McPherson et al., 2004), kappa can be used to compare model performance when species prevalence to train and test the models is the

same. Amongst the four SDMs used in this study, the BRT model had a significantly higher kappa value ($p = 0.05$; Table 2).

$$\varkappa = \frac{(TP + TN) - \left(\frac{(TP + FN) * (TP + FP) + (FP + TN) * (FN + TN)}{N}\right)}{N - \left(\frac{(TP + FN) * (TP + FP) + (FP + TN) * (FN + TN)}{N}\right)} \quad (1)$$

Threshold-independent metrics to assess SDM performance are preferred over threshold-dependent metrics, because a threshold does not need to be chosen by the modeler (Fielding & Bell, 1997; Franklin, 2010). However, since the use of k-fold cross-validation to test SDMs overinflates AUC scores due to spatial sorting bias, a new test dataset with spatial sorting bias removed was created as proposed by Hijmans (2012). The presence of spatial sorting bias is determined by looking at the ratio of Euclidian distances between presences and background points used to train and test the SDM. A ratio close to 0 indicates high spatial sorting bias while a value close to 1 indicates no spatial sorting bias is affecting model validation. Prior to this procedure when AUC scores were calculated in the traditional manner, the average geographic distance between training-presence/testing-presence cells, as well as training-presence/testing-background cells was approximately 10.8 and 301.7 km, respectively. Since the ratio between the presence cells and training-presence/testing-background cells was nearly zero ($0.036 = 10.8/301.7$), the initial model validation suffered from high spatial sorting bias (Hijmans, 2012). Pairwise distance sampling of the training dataset was performed using the "pwdSample" function in the dismo package (Hijmans et al., 2014), identifying 232 test points (116 presences and 116 background points). The average geographic distances for these test points was 317.1 and 296.5 km, respectively, thus removing spatial sorting bias since the ratio between them was approximately 1

(1.06 = 317.1/296.5). This validation procedure significantly changed the ranking of the SDMs (Table 2). Only the two tree-based SDMs, BRT and RF, retained high cAUC scores while the scores for MaxEnt and the GAM model dropped significantly ($p = 0.05$).

Current Rosy-Finch Distributions

Rosy-Finch distribution maps were created using Generic Mapping Tools 5.1.1 (Wessel, Smith, Scharroo, Luis, & Wobbe, 2013). Current distribution maps were created for each SDM as continuous probability surfaces (Figure 3). Grids were next reclassified to binary, presence-absence maps using a threshold maximizing the kappa statistic identified using the "PresenceAbsence" package (Figure 4; Freeman & Moisen, 2008b). The threshold maximizing kappa was used because Freeman and Moisen (2008a) found it to be the optimal threshold for rare species with low prevalence in their study comparing 11 different thresholds (threshold maximizing kappa was 0.36 for the BRT, 0.42 for the GAM, 0.44 for MaxEnt, and 0.34 for RF). The last map made for Rosy-Finch current distribution shows agreement of the four SDM presence-absence maps (Figure 5).

Future Rosy-Finch Distributions

Projected distributions for 2061–2080 were made using only the BRT and RF models since they had statistically significantly higher cAUC values than either the MaxEnt or GAM model ($p = 0.05$). Each GCM projection was first converted to a presence-absence map using maximum kappa as the threshold so that GCM agreement maps for RCP2.6 (15 GCMs) and RCP8.5 (17 GCMs) for the BRT and RF models, respectively, could be made (Figure 6). Consensus maps were then created by taking the mean of the projections made using the different GCMs from each emissions scenario to provide maps of continuous probability surfaces (Figure 7).

To determine the mean projected loss, as well as the variability between GCMs projecting loss of Rosy-Finch habitat in North America, change in suitable habitat was calculated using the "r.report" function (GRASS Development Team, 2012). The BRT and RF models were projected to each GCM (see Table 1) and converted to a binary presence-absence map using the threshold that maximized kappa (BRT= 0.36; RF= 0.34). Next, the difference was taken between current presence-absence and each presence-absence map from the different GCMs (Figure 8). The mean projected loss of suitable habitat (i.e., cells no longer classified as present in a presence-absence map) for the RCP2.6 emissions scenario was -42,620 km^2 95% CI [-69557, -15683] and -119,318 km^2 95% CI [-141275, -97361] for the BRT and RF models, respectively. The mean projected loss for the RCP8.5 emissions scenario was -146,379 km^2 95% CI [-187602, -105157] and -144,204 km^2 95% CI [-167541, -120867] for the BRT and RF models, respectively. Change in habitat suitability is depicted by showing areas where new suitable habitat is projected to occur, areas where suitable habitat is projected to be lost, and areas where no change in suitability is projected (Figure 9). This was calculated by taking the mode of individual grid cells from all GCM projections (i.e., 15 GCMs for RCP2.6 and 17 GCMs for RCP8.5), and taking the difference between it and the modeled current presence-absence map.

Variable Importance

Variable importance was determined for the two SDMs used to make projections: BRT and RF. Boosted regression trees determine variable importance by assessing the improvement in squared error at nonterminal splits made by each variable, summing the improvement from a single tree, and averaging this improvement over all trees used to fit the model (Friedman, 2001). The relative influence is calculated by normalizing the improvement due to each variable so that the contributions sum to 100 (Friedman, 2001). Cliffs were still an important

predictor, but annual precipitation had the highest relative influence in the BRT model (Table 3).

Random forest regression trees calculate prediction error for each tree using the out-of-bag observations (one-third of the dataset is randomly withheld when constructing a tree). For RF (as opposed to random forest classification trees), prediction error is recorded as mean squared error (Breiman, 2001). In order to calculate variable importance, values of each variable are permuted one at a time, and prediction error is calculated using the out-of-bag observations. The difference in the prediction errors is recorded over all trees, averaged, and standardized by dividing it by the standard deviation of the difference (Breiman, 2001; Liaw & Wiener, 2002). Permuting the values of one variable at a time simulates noise in that predictor, allowing its influence to the model's predictive accuracy to be quantified. Variable importance is reported as the percentage increase in mean squared error when the variable is removed (i.e., when values of it are permuted) from the model. Cliffs and mean annual temperature were the two most important variables to the RF model, and surprisingly, the biome variable had the least predictive skill (Table 4). Rosy-Finch breeding ecology occurs solely in the alpine-tundra biome and nests have never been outside of this biome (R. Johnson, personal communication, February 25, 2012). However, Rosy-Finches require the presence of cliffs or steep talus slopes to be present in this biome for it to be suitable. Since SDMs are highly influenced by the geographic extent used for training and testing and large expanses of Arctic tundra exist without a cliff component, this severely reduced the predictive skill of this variable.

Table 2

Species distribution model accuracy

Model	% Presences	% Background	Kappa	AUC	cAUC
BRT	0.943 (0.934–0.951)	0.997 (0.996–0.998)	0.956 (0.950–0.962)	0.998 (0.998–0.999)	0.972 (0.949–0.988)
GAM	0.757 (0.740–0.774)	0.981 (0.978–0.984)	0.787 (0.772–0.801)	0.980 (0.977–0.982)	0.776 (0.713–0.830)
MaxEnt	0.698 (0.679–0.715)	0.984 (0.982–0.987)	0.746 (0.731–0.762)	0.976 (0.973–0.978)	0.772 (0.711–0.828)
RF	0.767 (0.750–0.783)	0.985 (0.982–0.987)	0.802 (0.788–0.814)	0.980 (0.977–0.982)	0.975 (0.956–0.990)

Note. Accuracy assessed using both threshold-dependent and threshold-independent metrics. Reported values include the median from 10,000 bootstrap resamples and bootstrapped 95% CIs. Pairwise distance sampling provided 232 test points (116 presences and 116 background points) to calculate the cAUC. % Presences = presences predicted correctly using a threshold of 0.5; % Background = background points predicted correctly using a threshold of 0.5; Kappa was calculated using a threshold of 0.5; AUC = area under the curve of the receiver operating characteristic; cAUC = calibrated area under the curve of the receiver operating characteristic; BRT = boosted regression trees; GAM = generalized additive model; MaxEnt = maximum entropy; RF = random forest regression trees.

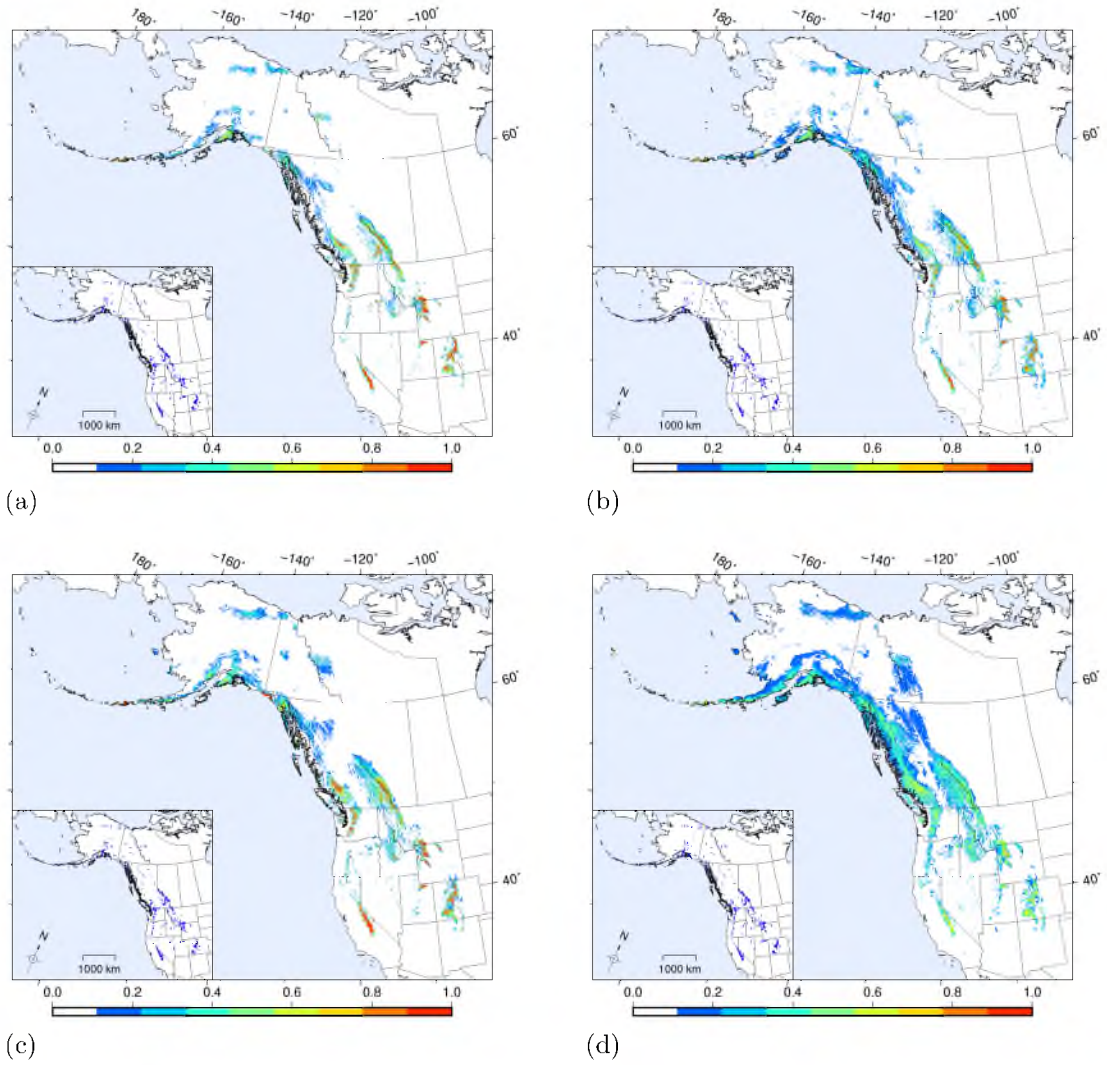


Figure 3. Rosy-Finch current geographic distribution maps. (a) BRT; (b) RF; (c) GAM; (d) MaxEnt. Probabilities below 0.06 and areas beyond the study extent are white. Inset map shows presences used to train and test SDMs. BRT = boosted regression trees; RF = random forest regression trees; GAM = generalized additive model; MaxEnt = maximum entropy; SDM = species distribution model.

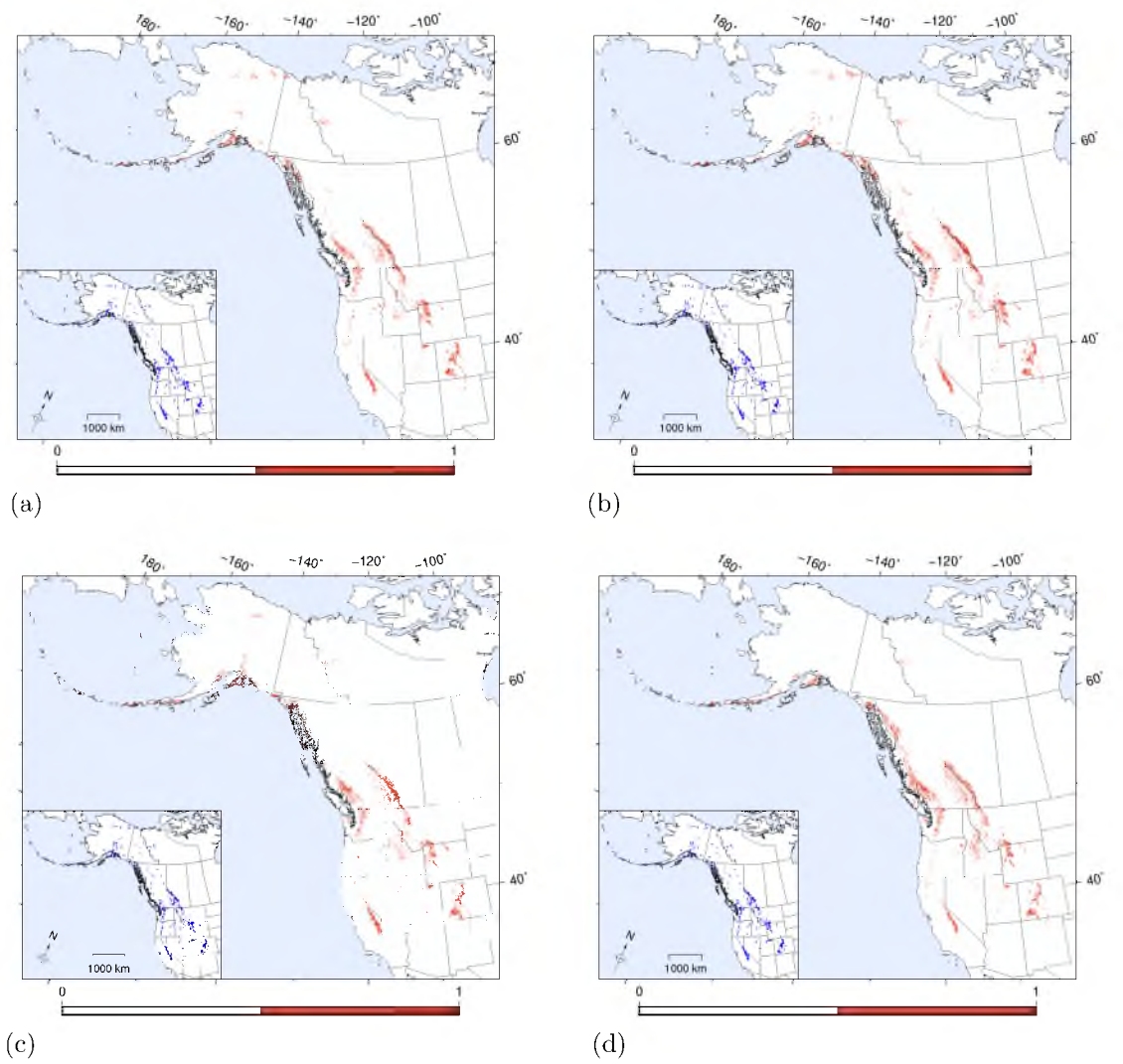


Figure 4. Rosy-Finch current presence-absence maps. (a) BRT; (b) RF; (c) GAM; (d) MaxEnt. The threshold maximizing kappa was used to create binary maps (BRT = 0.36; RF = 0.34; GAM = 0.44; MaxEnt = 0.44). Inset map shows presences used to train and test SDMs. BRT = boosted regression trees; RF = random forest regression trees; GAM = generalized additive model; MaxEnt = maximum entropy; SDM = species distribution model.

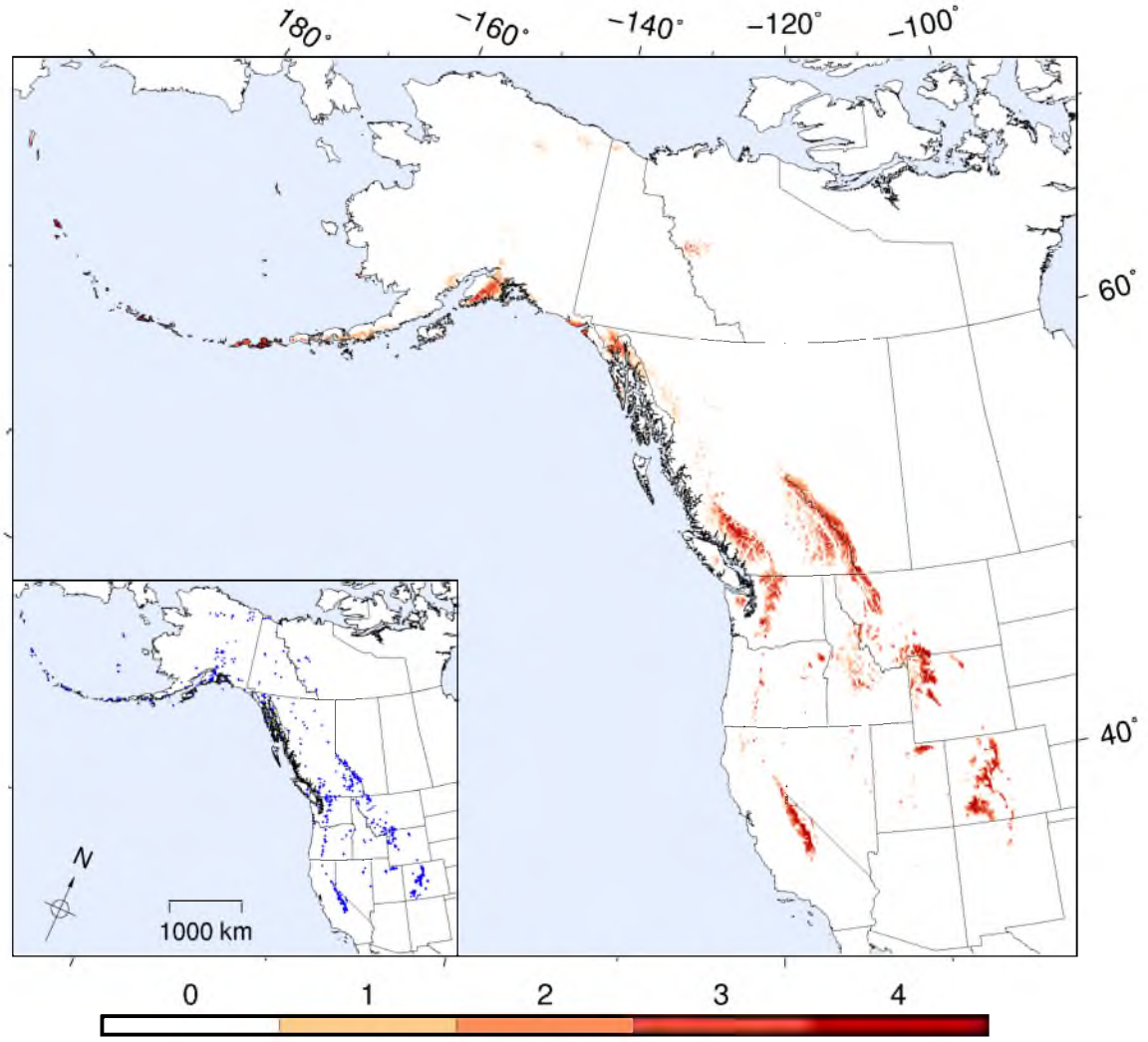


Figure 5. Rosy-Finch current agreement map. Each SDM was converted to a presence-absence map using a threshold maximizing kappa (BRT = 0.36; RF = 0.34; GAM = 0.44; MaxEnt = 0.44) and then the number of SDMs predicting presence for that grid cell were summed (e.g., a value of 4 indicates that all SDMs predict presence for a particular grid cell). Inset map shows presences used to train and test SDMs. BRT = boosted regression trees; RF = random forest regression trees; GAM = generalized additive model; MaxEnt = maximum entropy; SDM = species distribution model.

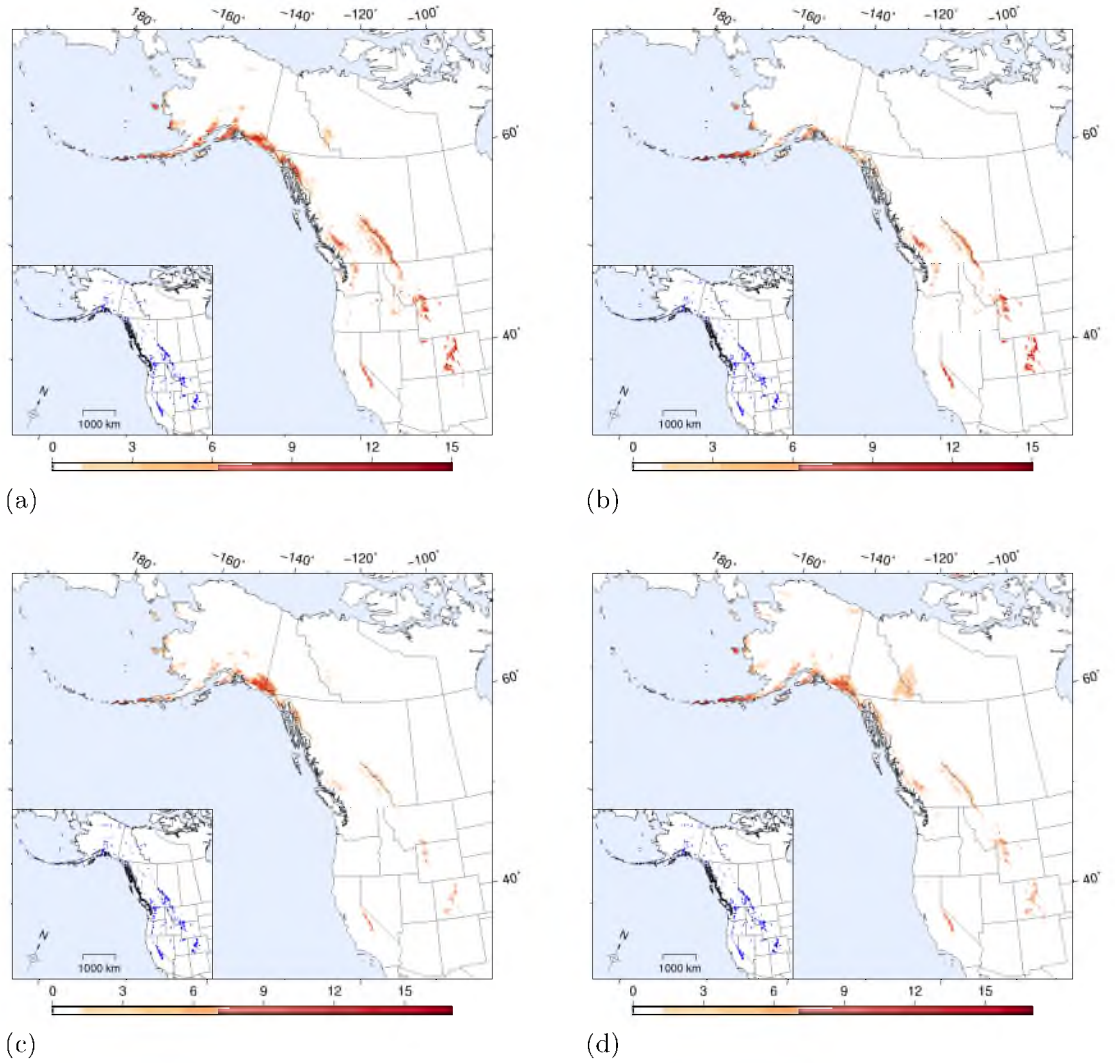


Figure 6. Rosy-Finch GCM agreement maps for 2061–2080. (a) BRT projection for RCP2.6; (b) RF projection for RCP2.6; (c) BRT projection for RCP8.5; (d) RF projection for RCP8.5. Values indicate the number of GCMs predicting presence for a given cell where presence was determined using a threshold that maximized kappa (BRT = 0.36; RF = 0.34). A total of 15 GCMs were available for RCP2.6 and 17 GCMs for RCP8.5. Inset map shows presences used to train and test SDMs. BRT = boosted regression trees; RF = random forest regression trees; GCM = general circulation model; RCP = representative concentration pathway; SDM = species distribution model.

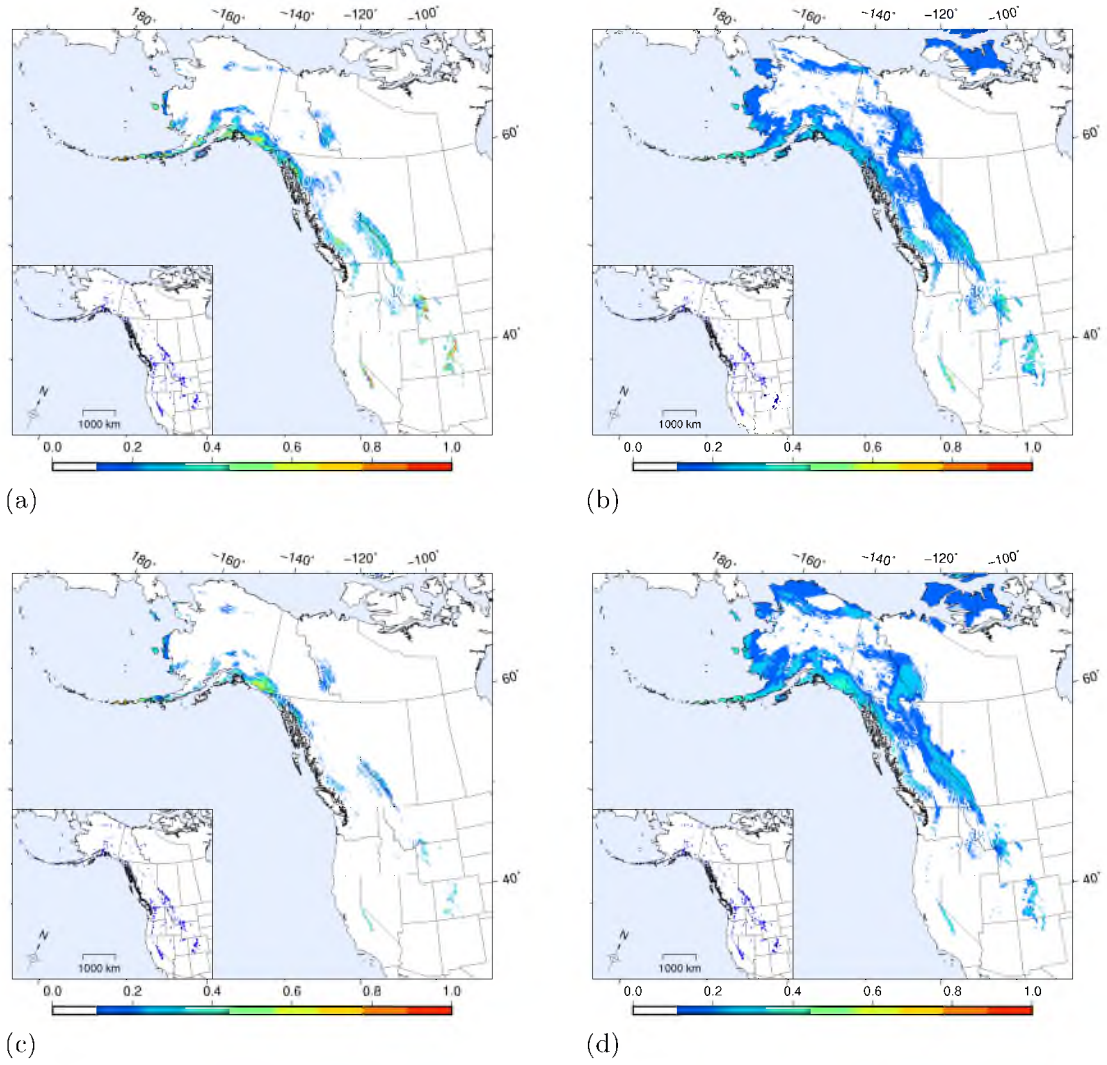
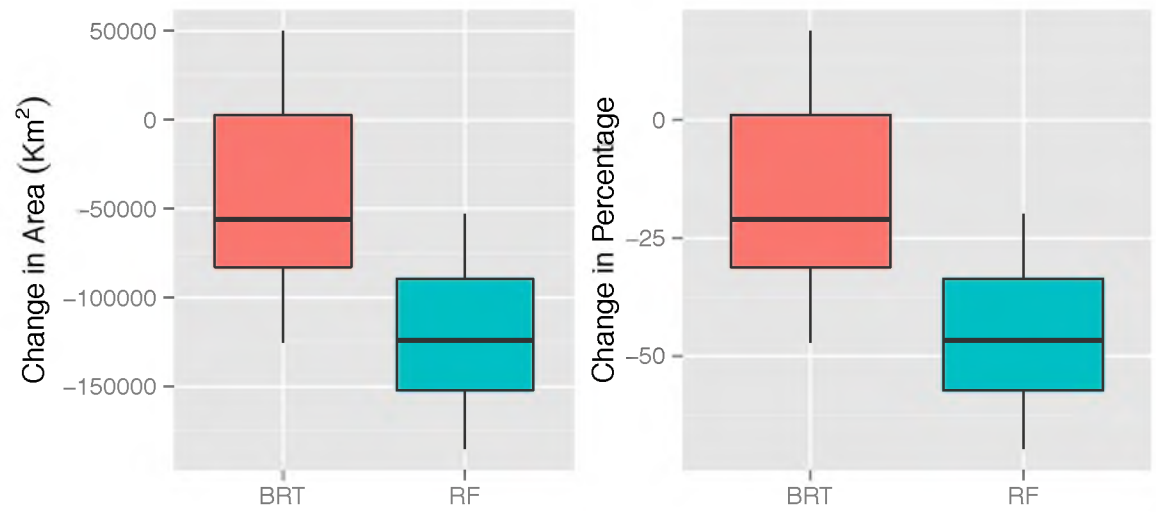
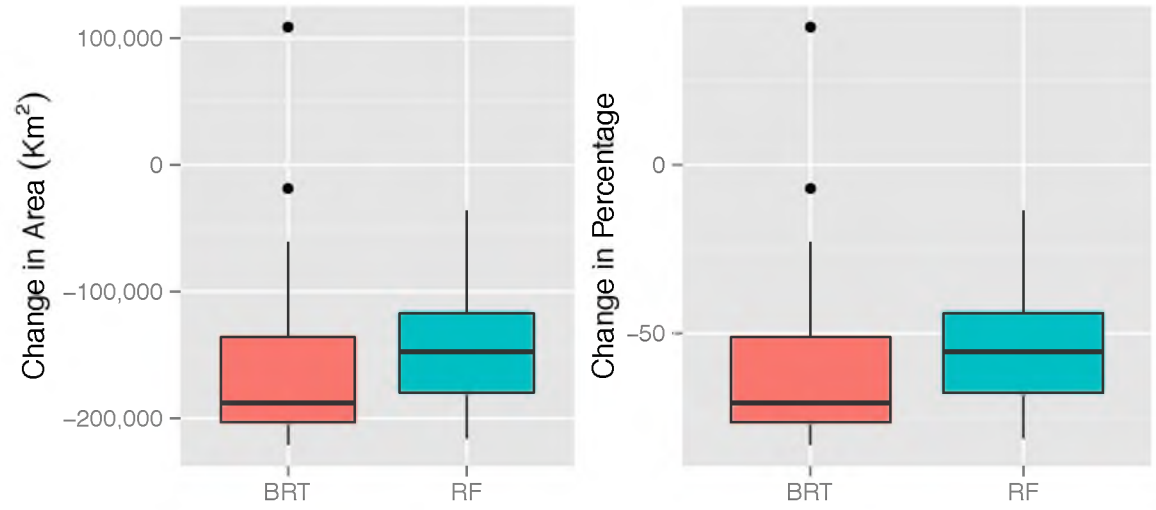


Figure 7. Rosy-Finch GCM consensus maps in 2061–2080. (a) BRT consensus projection for RCP2.6; (b) RF consensus projection for RCP2.6; (c) BRT consensus projection for RCP8.5; (d) RF consensus projection for RCP8.5. Statistical consensus was determined by taking the mean of the probabilities (i.e., taking the mean by individual grid cell) from 15 GCMs (RCP2.6) or 17 GCMs (RCP8.5). Probabilities below 0.06 and areas beyond the study extent are white. Inset map shows presences used to train and test SDMs. BRT = boosted regression trees; RF = random forest regression trees; GCM = general circulation model; RCP = representative concentration pathway; SDM = species distribution model.



(a)



(b)

Figure 8. Projected change: Area and percentage for 2061–2080. (a) BRT and RF projected change for RCP2.6 emissions scenario; (b) BRT and RF projected change for RCP8.5 emissions scenario. Change in area and percentage of suitable habitat calculated by taking the difference from each projected presence-absence map (15 GCMs for RCP2.6 and 17 GCMs for RCP8.5) and the current presence-absence maps. Binary presence-absence maps for both current and future projections were made using the threshold that maximized the kappa statistic (BRT= 0.36; RF= 0.34). The black line indicates the median loss and the box delineates the 25th and 75th percentiles of projected loss in area and percentage of suitable habitat. Whiskers are 1.5 $*$ IQR and outliers are demarcated as black dots. IQR = interquartile range; BRT = boosted regression trees; RF = random forest regression trees; GCM = general circulation model; RCP = representative concentration pathway.

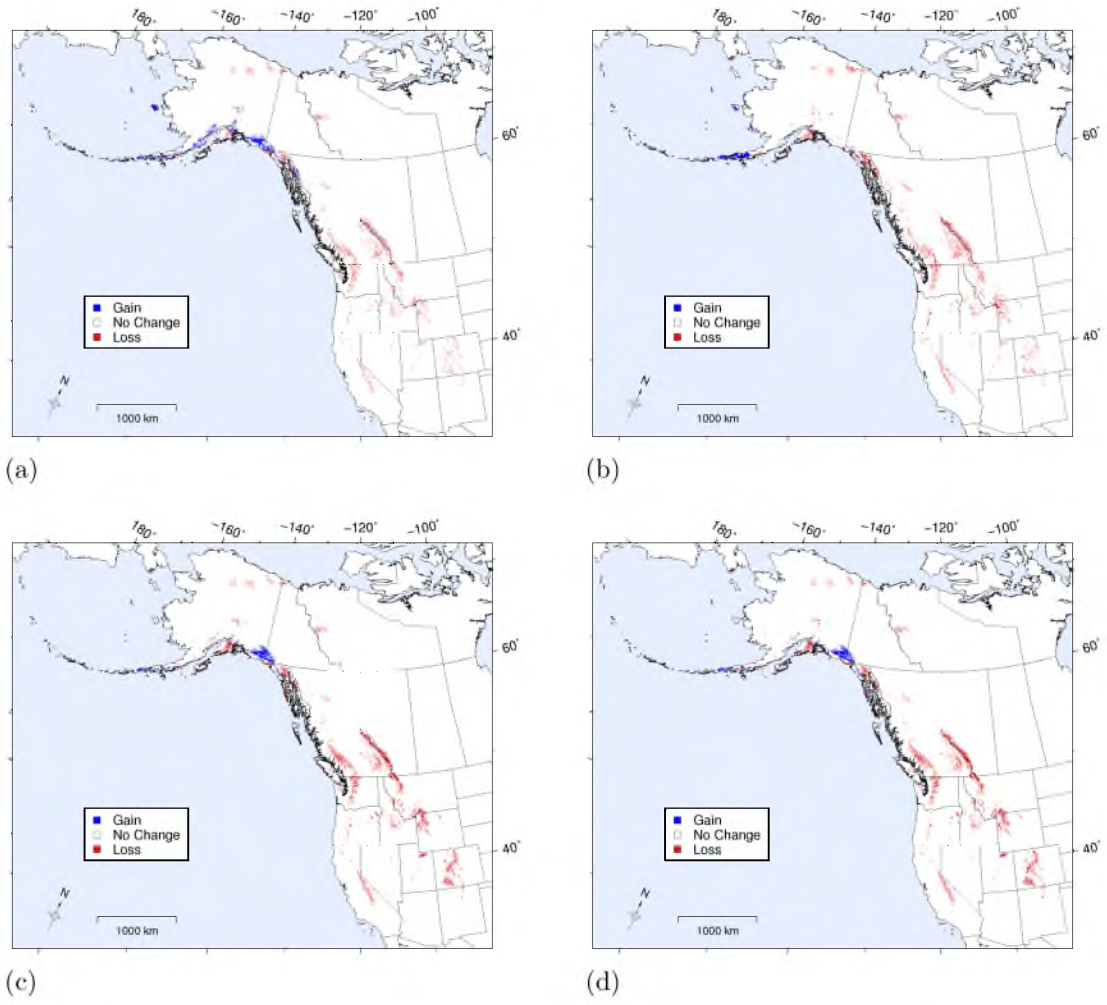


Figure 9. Projected change: Gains, losses, and no change for 2061 2080. (a) BRT projection for RCP2.6; (b) RF projection for RCP2.6; (c) BRT projection for RCP8.5; (d) RF projection for RCP8.5. Projections made with all the GCMs were converted to presence-absence maps using the threshold that maximized the kappa statistic (BRT= 0.36; RF= 0.34). Next, gains, losses, and no change were calculated by taking the mode of individual grid cells from the GCM projections and subtracting the current presence-absence map (also calculated using the threshold that maximized kappa). This difference determined whether an individual grid cell became suitable, lost suitability, or whether no change occurred. BRT = boosted regression trees; RF = random forest regression trees; GCM = general circulation model; RCP = representative concentration pathway.

Table 3

Boosted regression tree variable importance

Variable	Relative Influence
Annual Precipitation	22.952
Annual Mean Temperature	9.968
Isothermality	8.430
Cliffs	8.089
Temperature Seasonality	6.205
Mean Diurnal Range	5.299
Precipitation Seasonality	4.487
Max Temperature of Warmest Month	4.397
Biome	4.356
Precipitation of Driest Month	3.456
Precipitation of Warmest Quarter	3.409
Mean Temperature of Wettest Quarter	3.394
Mean Temperature of Warmest Quarter	2.867
Min Temperature of Coldest Month	2.539
Precipitation of Coldest Quarter	2.507
Precipitation of Driest Quarter	2.380
Mean Temperature of Coldest Quarter	1.953
Precipitation of Wettest Quarter	1.689
Precipitation of Wettest Month	1.625

Note. Relative influence of predictor variables in a BRT model fitted by 10-fold cross-validation. Model parameters: tree complexity = 10, learning rate = 0.01, bagging fraction = 0.6, number of trees = 3450. The variables "temperature annual range" and "mean temperature of driest quarter" had zero influence and were not used for final BRT model. BRT = boosted regression trees.

Table 4

Random forest regression trees variable importance

Variable	% Increase MSE
Cliffs	52.948
Annual Mean Temperature	51.590
Precipitation Seasonality	49.300
Isothermality	48.501
Mean Diurnal Range	45.973
Temperature Seasonality	45.862
Mean Temperature of Wettest Quarter	45.201
Precipitation of Coldest Quarter	39.238
Precipitation of Driest Month	38.756
Annual Precipitation	37.955
Min Temperature of Coldest Month	35.146
Precipitation of Driest Quarter	34.162
Mean Temperature of Warmest Quarter	34.148
Mean Temperature of Driest Quarter	33.664
Mean Temperature of Coldest Quarter	31.971
Max Temperature of Warmest Month	30.986
Temperature Annual Range	30.805
Precipitation of Wettest Quarter	30.521
Precipitation of Warmest Quarter	30.004
Precipitation of Wettest Month	28.060
Biome	22.496

Note. Variable importance in a RF model fitted by 10-fold cross-validation. Model parameters: predictors available at each split = 7, number of trees = 1000. MSE = mean squared error; RF = random forest regression trees.

DISCUSSION

The results from this analysis suggest that Rosy-Finches are extremely vulnerable to projected climate change in the next half-century, particularly populations in northern Utah where projected warming is the highest in the contiguous United States according to GCMs used in this analysis (GCMs used in CMIP5 and the IPCC AR5). Rosy-Finches occupy a narrow ecological niche, collectively breeding at a higher elevation than any other North American breeding bird (>700 species north of the United States-Mexico border) from midlatitudes to polar regions in western North America. Their specific breeding requirements are exclusively cliffs or steep talus slopes as nest substrate for avoidance of land predators, and nearby vegetation belonging to the alpine-tundra biome for suitable foraging habitat. Today, such habitat is already restricted to the highest mountains in the southerly portions of their range and much of this area is projected to disappear as biomes shift (Williams, Jackson, & Kutzbach, 2007), preventing an opportunity for Rosy-Finches to track climate warming.

Evidence does exist that treelines have advanced upslope as predicted due to rising temperatures (e.g., Klanderud & Birks, 2003; Kullman, 2002). Treelines are largely thermally controlled (Körner, 1998), but their exact demarcation is also due to nutrient availability. In addition to warmer temperatures providing more optimal growing conditions at higher elevations, warmer temperatures will also increase nitrogen and phosphorous availability to plants due to increased mineralization (Chapin III, Shaver, Giblin, Nadelhoffer, & Laundre, 1995). Warming experiments provide evidence that vegetation in tundra ecosystems responds rapidly to 1–3°C temperature increases, thereby providing robust experimental evidence that recent

shrub encroachment into these ecosystems is due to climate warming (Walker et al., 2006).

High-elevation species are disproportionately threatened by climate warming because poleward and upward shifts of ecosystems will reduce habitat for species with distributions restricted to narrow elevational boundaries and higher temperatures can increase metabolic costs for specialized species (Dirnböck, Essl, & Rabitsch, 2011; Sekercioglu, Schneider, Fay, & Loarie, 2008). Furthermore, new research suggests that mountains are warming faster than lower elevations due to elevation-dependent warming caused in-part by factors related to albedo, clouds, water vapor and radiative fluxes, and aerosols (Mountain Research Initiative EDW Working Group, 2015). This is significant because elevation-dependent warming is poorly understood and is not yet adequately incorporated into general circulation models meaning that projected warming in high elevations could be underestimated (Mountain Research Initiative EDW Working Group, 2015). A well-studied example of a mammal species in North America inhabiting montane habitats threatened by climate warming is the American Pika (*Ochotona princeps*; Beever, Brussard, and Berger 2003). Experimental research has shown that the American Pika has a high body temperature ($\bar{x}$ = 40.1°C, range 37.9–42.7°C) and that exposure to ambient temperatures as low as 25.5–29.5°C are lethal (MacArthur & Wang, 1973; Smith, 1974). Recently, American Pikas have been discovered to be extirpated from nine historically occupied sites across the Great Basin — sites that were characterized as being hotter and drier and occurring at lower elevations compared to sites where pikas continue to persist (Beever et al., 2003; Beever, Ray, Mote, & Wilkening, 2010). Although a few low-elevation American Pika populations persist where suitable talus fields exist, for instance at the Columbia River Gorge in Oregon (Simpson, 2009) and Hays Canyon in Nevada (Beever, Wilkening, McIvor, Weber, & Brussard, 2008), their above-ground behavior is crepuscular rather than diurnal as

occurs at high-elevation sites (Smith, 1974). Low-elevation American Pika populations deserve continued research and monitoring, but their continued existence suggests that a behavioral response together with an available refugium underneath talus slopes, allows these populations to persist at present. The projected distributions of Rosy-Finches strongly suggest that their fate is similarly linked to climate warming; however, unlike the American Pika, their upper thermal tolerance limits are unknown. Uncertainty remains whether the risk climate warming poses for Rosy-Finches is due to higher projected temperatures exceeding their upper thermal tolerances, or due to the loss of available foraging habitat and/or changes in abundance, composition, and species of their food supply as ecosystems respond to climate warming.

Limitations of Species Distribution Models

Species distribution models make statistical relationships between geographic locations where an organism is found and the environmental values occurring there in relation to the values of the variables across the landscape to project habitat suitability onto a map. The complexity of the response functions used to make these correlations is determined by the number of presences available for training and testing the SDM, as well as the actual SDM algorithm. Although BRT and RF can fit complex response functions with abrupt thresholds (a beneficial quality due to the Rosy-Finch's specific breeding requirements: cliffs vs. no cliffs; alpine-tundra vs. other biomes), they do not model an organism's ecological niche mechanistically since ecophysiological relationships are not included in the modeling process. The optimal modeling approach according to Morin and Thuiller (2009) would be to compare SDM predictions to a process-based model to determine areas of model agreement between correlative and mechanistic approaches, but at present, insufficient data exist to carry this out.

Species distribution models are often criticized because they fail to include

biotic interactions. Biotic interactions include consumed resources directly impacted by a species (e.g., soil nutrients), interspecific biotic interactions (e.g., competition, predation, facilitation, parasitism), or even describe a species' dispersal capabilities (Pearson & Dawson, 2003). Peer-reviewed articles that have used biotic interactions to improve SDM predictions at coarse-grain scales are uncommon in the SDM literature. Araújo and Luoto's (2007) paper was the first to show that the inclusion of a biotic interaction (facilitation) — the inclusion of a host plant important to a species of butterfly in Europe — improved SDM predictions of climate change impacts at macroecological scales. Similarly, Heikkinen, Luoto, Virkkala, Pearson, and Körber (2007) using Araújo and Luoto's (2007) methodological approach, showed that including a biotic interaction (again facilitation) — the inclusion of woodpecker species responsible for excavating cavities used by small owls — improved SDM accuracy for current owl distributions at a 10 km grain, and somewhat improved SDM accuracy at a 40 km grain in Scandinavia. However, the inclusion of biotic interactions has yet to be shown to improve predictive skill at finer resolutions (this analysis has a resolution of 828 m^2). The inclusion of biotic interactions is less critical in modeling the ecological niche of the Rosy-Finch since its ecological niche can be well quantified by available environmental grids. In the case of modeling Rosy-Finches, having higher resolution climatic and vegetation data is more important than biotic interactions for improving SDM accuracy. Higher resolution grids would better capture the complex topography of western North America allowing suitable sky islands to be identified. This would improve statistical relationships between the presence data and environmental grids and thus improve predictive performance. For instance, in this study, some records with high geographic precision that were located in the alpine-tundra biome, were assigned values representative of biomes downslope simply due to the coarseness of the 828m^2 input grids in regions characterized by complex topography.

Another valid criticism made of SDMs is that projected warming will bring with it nonanalogue climates, whereby climatic values beyond the range of values used to train correlative models will occur. Making extrapolations in time or space into nonanalogue climates is risky and such predictions face high uncertainties (Fitzpatrick & Hargrove, 2009). However, in the case of this study, nonanalogue climate is less of a concern since geographic regions with projected nonanalogue climate occur predominately towards the equator, well beyond the distribution of this alpine obligate (Williams & Jackson, 2007; Williams et al., 2007). As the results from this study and analyses making projections of biome shifts show (e.g., Rehfeldt et al., 2012; Williams et al., 2007), Rosy-Finches are extremely vulnerable to climate change due to the projected disappearance of alpine-tundra biome in large parts of their current geographic range.

Another limitation to SDMs is that it is difficult to determine whether or not all Rosy-Finches will be similarly affected by climate change. Partners in Flight (2013) estimates the current global population of Black Rosy-Finch to be 20,000, compared to 40,000 for Brown-capped Rosy-Finch and 200,000 for Gray-crowned Rosy-Finch. Furthermore, projected temperature increases made by the IPCC AR5 are highest in the contiguous United States in regions that overlay the current breeding range of the Black Rosy-Finch (Figure 10). A limitation of SDMs as implemented in this analysis is that the competitive biotic interactions that exist between Nearctic Rosy-Finches (if any at all) cannot be determined. Plumage differences between Rosy-Finch species show abrupt differences at current range boundaries, and do not show the clinal gradation as suggested by the principal component analysis of plumage done by Drovetski et al. (2009), which disregarded the most distinctive phenotypic trait supportive of the three separate species — the pattern of gray on the head (Johnson, 1973). An abrupt rather than clinal gradation in the phenotype between Nearctic Rosy-Finch species is suggestive that local

adaptation to the environment has occurred. However, variation in the phenotype (currently used to justify three separate Nearctic Rosy-Finch species) may also be explainable due to phenotypic plasticity and differences in how the Rosy-Finch genotype is expressed across different environments. The dialogue of the North American Classification Committee (Chesser et al., 2013) makes it clear that neither the research conducted by Johnson (1973) nor Drovetski et al. (2009) conclusively answers this evolutionary biology question, and the present research makes no attempt at all to disentangle Rosy-Finch systematics. A limitation of the present study is that while projecting shifts in distribution for the Rosy-Finch superspecies is possible, projecting shifts of the individual Rosy-Finch species with evolved local adaptations to specific environments is exceedingly difficult because SDMs are highly influenced by the spatial extent chosen for model training and testing. When used for projecting distributions in time and space, SDMs should be built using the entire spatial extent of the species' distribution in order to capture the full variability of the environmental space (Thuiller, Brotons, Araújo, & Lavorel, 2004). Creation of three separate SDMs, one each for the Black Rosy-Finch, Brown-capped Rosy-Finch, and Gray-crowned Rosy-Finch would be possible for current distributions by using convex hulls surrounding presence points to determine spatial extents of the environmental space. However, such an approach would be inadequate to make projections in space and time because projections must be made to the same spatial extent used to train and test the SDM; therefore, potential gains of suitable habitat outside of this restricted geographic extent could not be determined.

Implications for Management

The short-term implications of this study for wildlife management are the need for an immediate assessment of Rosy-Finches, particularly the Black Rosy-Finch, which has the lowest estimated population and is designated as a "Species of Greatest Conservation Need," by Utah's Wildlife Action Plan (Sutter et

al., 2005). The assessment should include surveys targeted at existing suitable habitats as predicted by a SDM to determine their population status. Surveys in Utah and elsewhere should revisit historically occupied Black Rosy-Finch breeding sites lacking records in eBird and GBIF to determine whether populations there have been extirpated due to climate warming, or whether there is simply a lack of reporting coverage. In Utah, these include the Raft River Range and La Sal and Tushar Mountains (Johnson, 2002). Surveys would provide initial population estimates, as well as provide field validation of SDM predictions, which would consequently provide greater confidence in predictions of Rosy-Finch's response to climate change.

The long-term implications of this study for land management are to consider the use of large herbivores in alpine and tundra ecosystems to mitigate ecosystem response to climate change (Dirnböck, Dullinger, & Grabherr, 2003; Post & Pedersen, 2008) in order to preserve foraging habitat critical to Rosy-Finches during their breeding season. Temperature and nutrient availability are not the sole factors determining treeline; and the role of herbivory as a strong selective pressure on vegetation in these ecosystems is now recognized (Cairns & Moen, 2004; Oksanen & Ranta, 1992). Virtanen et al. (2010) revisited sites that were initially sampled 26 years earlier along the forest-tundra ecotone and found that herbivory had buffered the treeline's response to climate warming at many sites. Exclosure experiments at small-scales (Olofsson et al., 2009; Post & Pedersen, 2008), as well as at landscape-scales (Speed, Austrheim, Hester, & Mysterud, 2010), provide robust experimental evidence that herbivory from large herbivores can limit the advancement of the treeline due to climate warming by reducing aboveground biomass and selectively nibbling on young saplings. From a management perspective, high-densities of large herbivores are not needed to mitigate treeline response to climate warming. Speed et al. (2010) showed that low densities of

domestic sheep (*Ovis aires*; 25 sheep / km^2) were nearly as effective at limiting the treeline as were high-densities (80 sheep/ km^2). Considering that sheep herders already graze their sheep in alpine areas such as the High Uinta Wilderness in Utah (Figure 11), new land management may not be required in all areas. However, careful consideration should be given as to whether domestic sheep grazing should be expanded, or allowed to continue at all in present areas. Dual management of domestic sheep and Bighorn Sheep (*Ovis canadensis*) must be avoided because Bighorn Sheep easily contract *Pasteurella* spp. strains (*P. haemolytica* recently reclassified as *Mannheimia haemolytica* and *P. trehalosi, and P. multocida*) when their ranges overlap and die from acute hemorrhagic pneumonia (Foreyt & Jessup, 1982; George, Martin, Lukacs, & Miller, 2008). Experimental enclosure studies placing Bighorn Sheep and domestic sheep together (Foreyt, 1989) and experimental inoculation of healthy Bighorn Sheep with *M. haemolytica* obtained from domestic sheep (Foreyt, Snipes, & Kasten, 1994) have confirmed high virulence of *M. haemolytica* in Bighorn Sheep and have shown that it causes high mortality rates. Researchers have advised that conservation efforts for Bighorn Sheep should focus on large patches of suitable habitat $\geq$ 23 km from domestic sheep (Singer, Zeigenfuss, & Spicer, 2001) in order to minimize the risk of die-offs, which are today's biggest threat to Bighorn Sheep conservation (Gross, Singer, & Moses, 2000). However, prior to considering implementing such management actions as part of Rosy-Finch conservation efforts, experimental research should be conducted on Rosy-Finch physiology to determine their thermal neutral zone. Ethical considerations preclude determining Rosy-Finch's upper thermal tolerance limit, but knowledge of their thermal neutral zone provides empirical data on whether or not they will incur metabolic costs due to climate warming. This would allow wildlife managers to know whether or not mitigative actions put into place to suppress the treeline from moving upslope will be effective at conserving Rosy-Finches in the long-term.

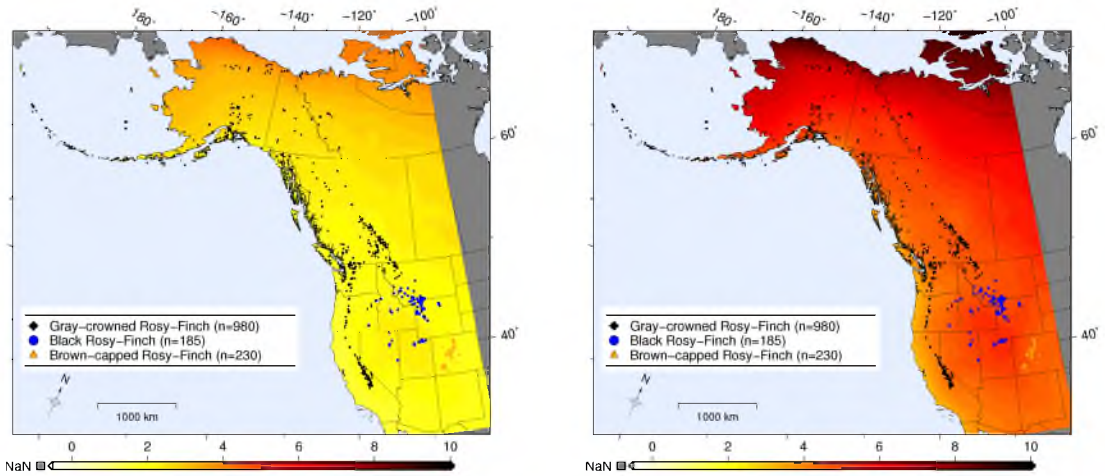


Figure 10. GCM consensus temperature anomalies (°C) for 2061–2080. (a) Anomaly for the RCP2.6 emissions scenario (mean of 15 GCMs); (b) Anomaly for the RCP8.5 emissions scenario (mean of 17 GCMs). Anomaly for mean annual temperature calculated by taking the difference of each GCM run from the 1961–1990 reference period and the 2061–2080 target period. The difference is statistically downscaled and applied to the 1950–2000 WorldClim mean annual temperature grid. Rosy-Finch presences overlaid. GCM = general circulation model; RCP = representative concentration pathway.

Figure 11. Herbivory mitigates ecosystem response to climate warming. Experimental research has shown that managing large herbivores above the treeline effectively prevents the treeline from moving upslope, thus protecting alpine-tundra ecosystems from tree and shrub encroachment. Domestic sheep graze above treeline in the High Uinta Wilderness in Utah. Photo by Veronica Elaine. Used with Permission.

CONCLUSION

In this study, uncertainties regarding choice of SDM, validation statistic, emissions scenario, and GCM were addressed in order to gain confidence in the predictions being made for Rosy-Finch distributions in 2061–2080. Since the ecological niche of the Rosy-Finch can be well captured using a SDM, this provided a unique opportunity to investigate the uncertainties regarding these other factors. The use of the cAUC with spatial sorting bias removed, provided a more robust measure for model validation, and allowed a ranking of SDMs to be made. Hijmans (2012, p. 686) stated the following:

> Comparisons of model results between species, regions, or any other results obtained with a different combination of model training and testing data are questionable if the results are not calibrated. The results presented here therefore question the conclusions of previous work that was based on comparing uncalibrated AUC values between different models.

Although this study did not compare SDMs across multiple species, overinflation of AUC scores does occur using k-fold cross-validation, and provides modelers with a false sense of their model's predictive skill. Perhaps the most interesting result of this study was that prior to using the cAUC, the BRT model was statistically better than the other three SDMs yet no statistically significant differences existed between the GAM, MaxEnt, and RF models ($p = 0.05$). After calculating the cAUC, the two tree-based approaches retained high scores, but the cAUC scores for the MaxEnt and the GAM model significantly dropped ($p = 0.05$). It is unclear why the BRT and RF models had higher predictive skill as measured by the cAUC, although it could be because both approaches have machine learning origins where predictive skill is emphasized using ensembles of trees rather than an emphasis on explanation.

An interesting question involving further research would be to create simulated data for an artificial species where each aspect of the niche is known precisely, and to create models that could be validated using the cAUC. Such a methodological approach has previously allowed modelers insights into their SDMs (e.g., Elith & Graham, 2009; Hirzel, Helfer, & Metral, 2001) and might provide an explanation for why the two approaches using ensembles of trees had higher cAUC scores.

This study also explored uncertainty in predictions in time by using multiple GCMs and emissions scenarios. Although I did not explore all four RCPs used by the IPCC AR5, this study used the low (RCP2.6) and high (RCP8.5) emissions scenarios, which provided an envelope for the range of expected responses to projected climate change dependent on long-term socio-economic, environmental, and technological, trends (Moss et al., 2010). As expected, projected losses of habitat were higher for the higher emissions scenario. Less clear, however, was why the RF model projected greater losses of suitable habitat for RCP2.6 than did the BRT model, yet the reverse was true for RCP8.5. Variability in predictions made using different GCMs has often been cited as a source of uncertainty for projecting distributions in space and time, yet rarely has this uncertainty been captured with statistical rigor. Porfirio et al. (2014) conducted a literature review in which they documented only seven studies using >10 GCMs to capture GCM uncertainty. Considering that SDMs are most often applied to predicting geographic distributions across space and time, and that making predictions is fraught with uncertainties, this is surprising. I provide an estimate of the loss of suitable habitat and quantify variability using 15 and 17 GCMs for the RCP2.6 and RCP8.5 emissions scenarios, respectively (18 of the 61 GCMs used by CMIP5). The BRT model had two outliers for the RCP8.5 emissions scenario; however, inspection of the predictions showed that predicted gains in suitable area occurred in Arctic tundra along the periphery of the northwestern Alaskan coastline in areas lacking suitable

nest substrate. This suggests that the climatic conditions from these GCMs became optimal in these geographic regions, but that the BRT model needs to weight cliffs higher, since the area remains unsuitable per Rosy-Finch breeding ecology.

This analysis followed a robust modeling framework for how climate change impacts analyses can be carried out using SDMs. Although uncertainty is an inherent consequence whenever predictions are being made, this analysis made significant strides in quantifying multiple sources of uncertainty. The results show that although variability exists in the predicted loss of suitable Rosy-Finch habitat, there is high agreement between the SDMs, GCMs, and emissions scenarios as to where Rosy-Finches will persist and the areas where they will become extirpated. This suggests that Rosy-Finches are extremely vulnerable to projected anthropogenic climate warming.

APPENDIX: FUTURE CLIMATE CONSENSUS GRIDS

Figure 12. Future climate consensus grids for RCP2.6. Consensus grids were created by taking the arithmetic mean of each variable across 15 GCMs. (a) Annual mean temperature; (b) Mean diurnal temperature range: mean of monthly (max temperature − min temperature); (c) Isothermality: mean diurnal range/temperature annual range; (d) Temperature seasonality: temperature standard deviation; (e) Max temperature of warmest month; (f) Minimum temperature of coldest month; (g) Temperature annual range: max temperature of warmest month − min temperature of coldest month; (h) Mean temperature of wettest quarter; (i) Mean temperature of driest quarter; (j) Mean temperature of warmest quarter; (k) Mean temperature of coldest quarter; (l) Annual precipitation in mm; (m) Precipitation of wettest month; (n) Precipitation of driest month; (o) Precipitation seasonality (coefficient of variation); (p) Precipitation of wettest quarter; (q) Precipitation of driest quarter; (r) Precipitation of warmest quarter; (s) Precipitation of coldest quarter; (t) Alpine-tundra biome; (u) Cliffs with 2 km buffer. GCM = global climate model; NaN = not a number (area beyond study extent with no data); RCP = representative concentration pathway.

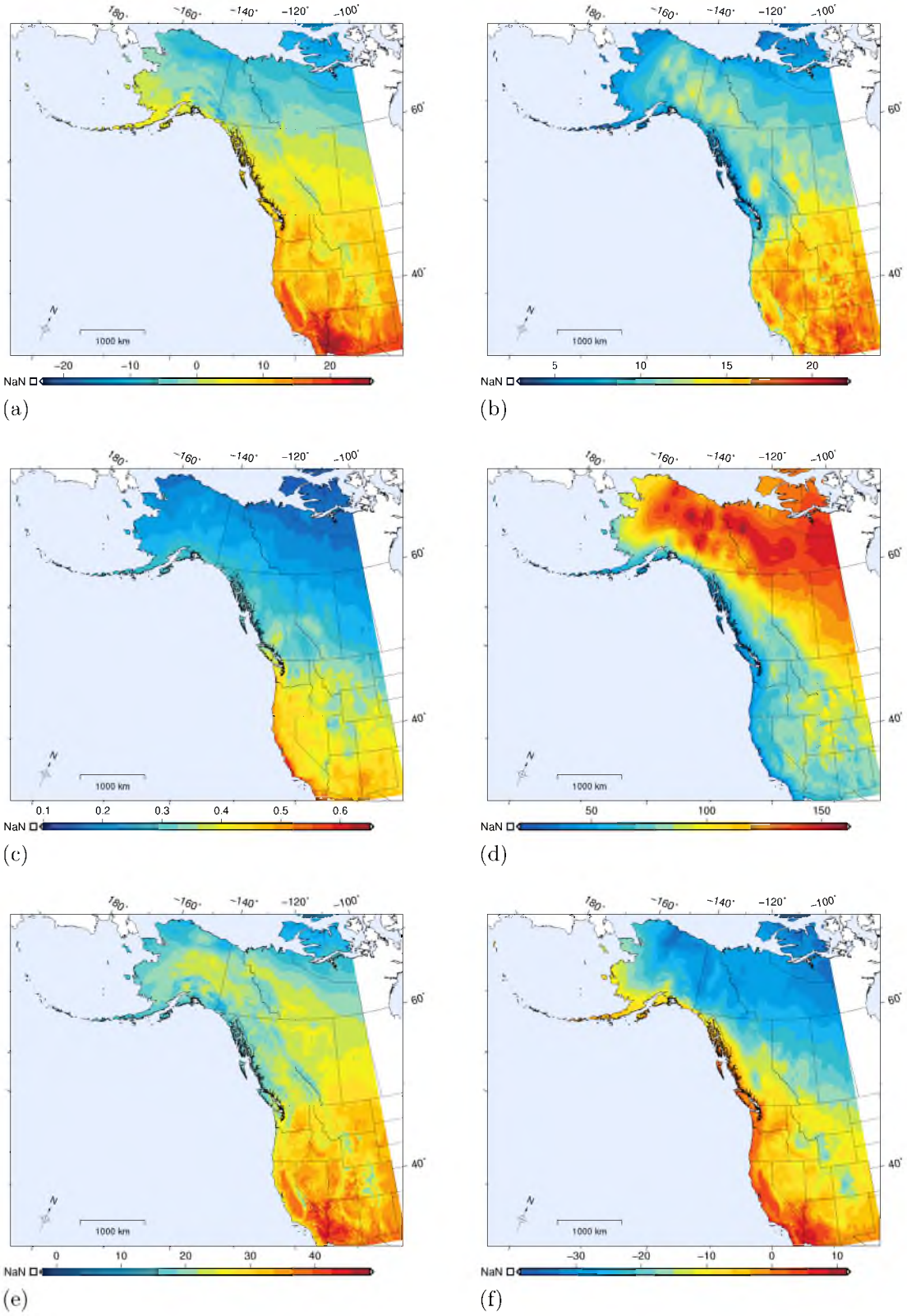


Figure 12. continued.

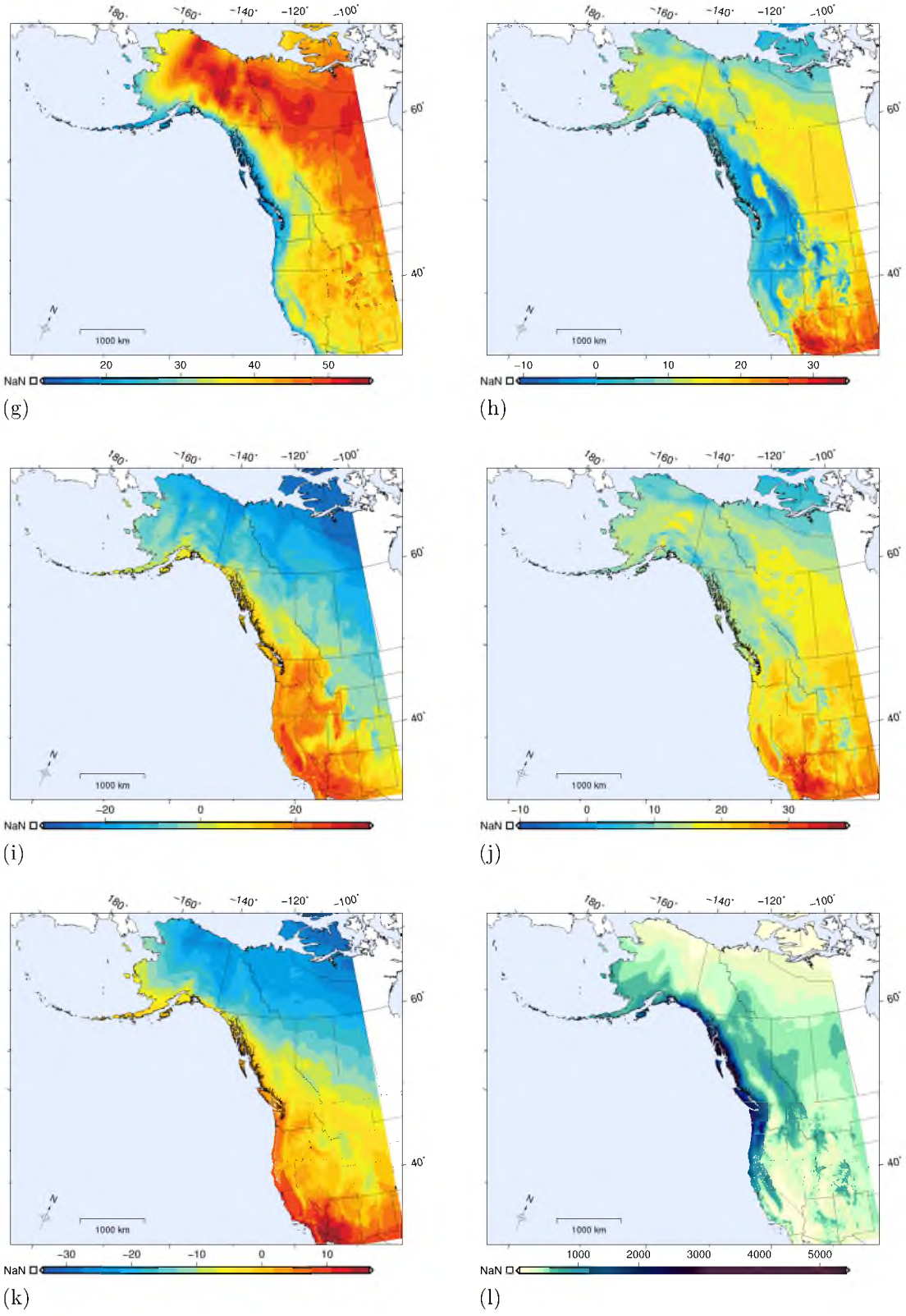


Figure 12. continued.

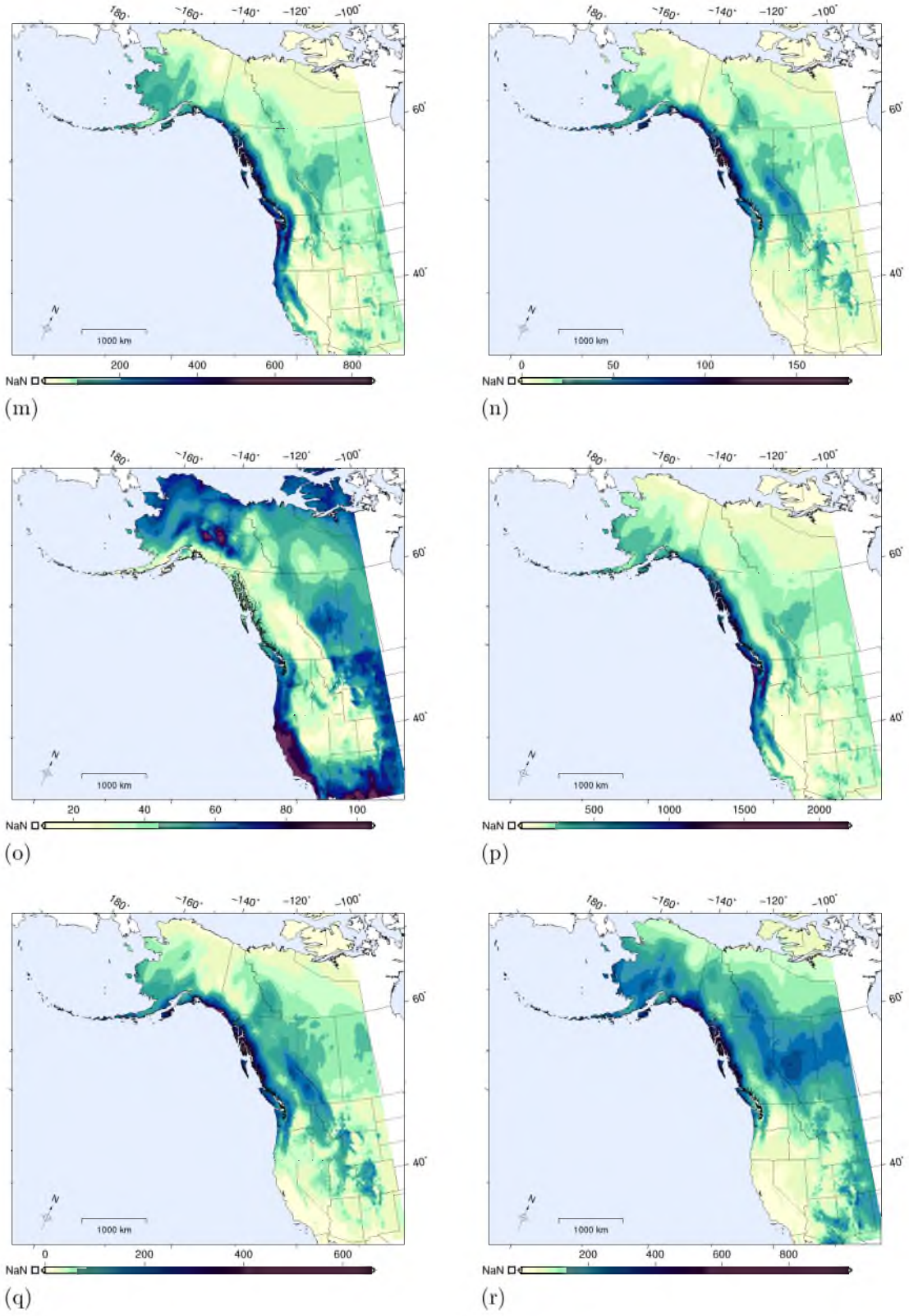
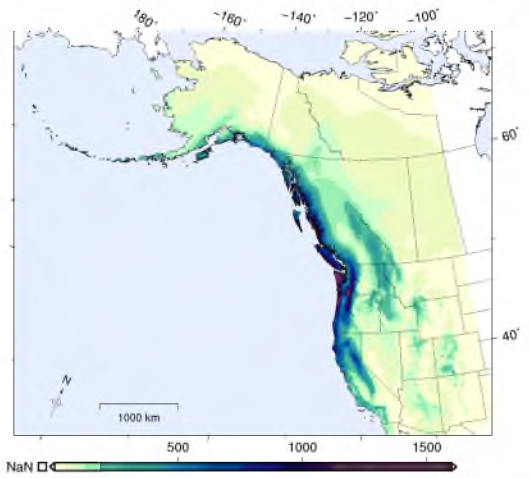
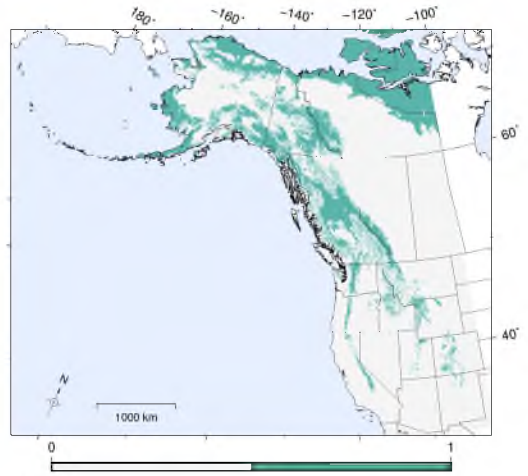


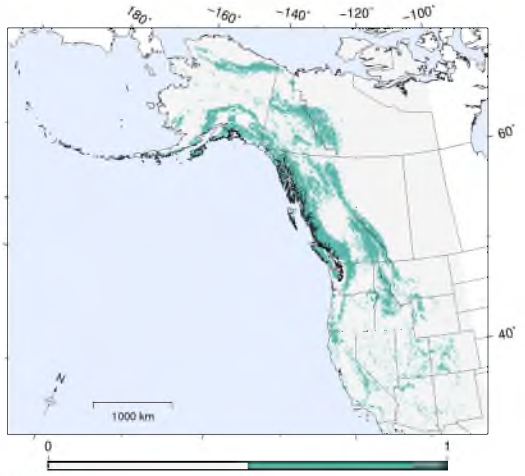
Figure 12. continued.



(s)



(t)



(u)

Figure 12. continued.

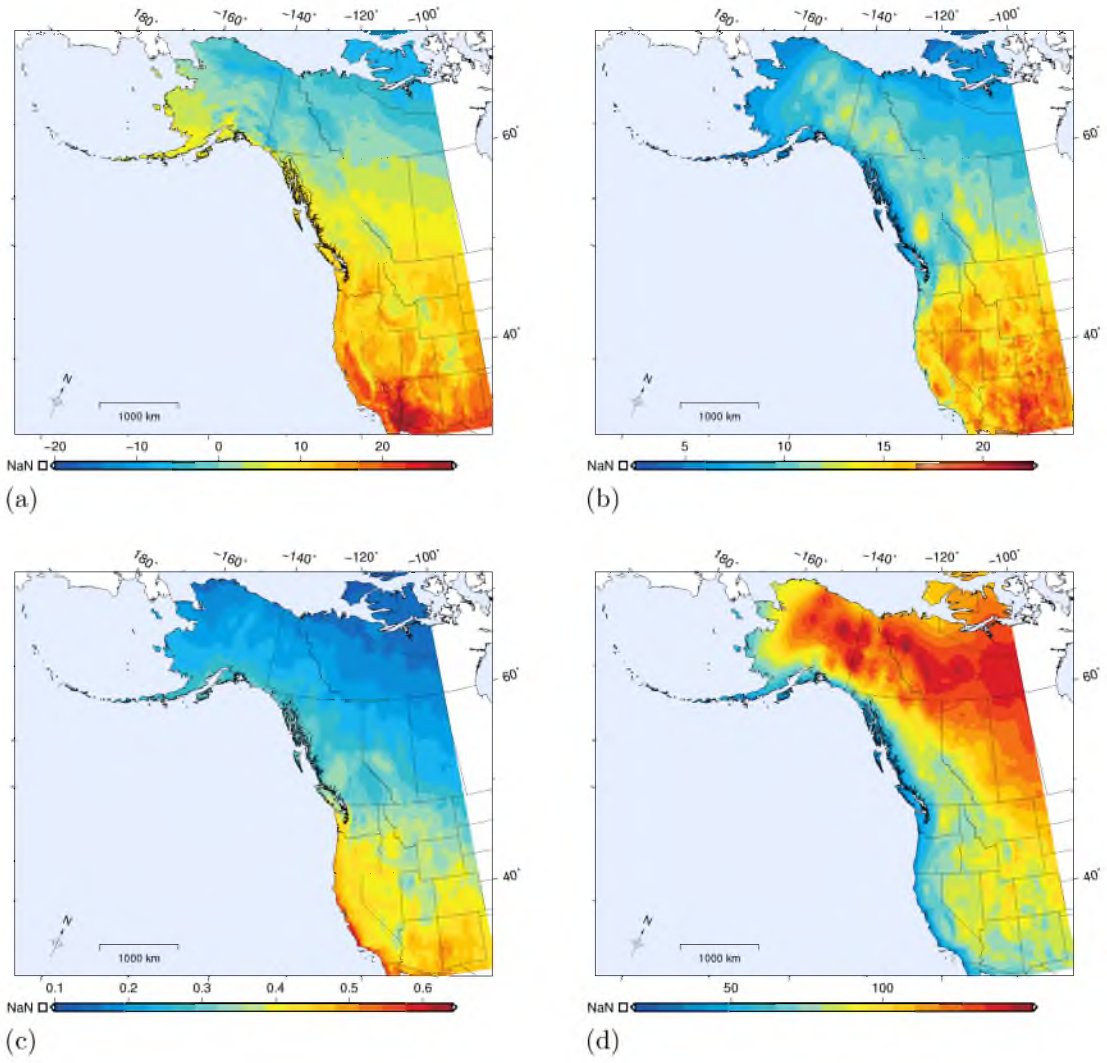


Figure 13. Future climate consensus grids for RCP8.5. Consensus grids were created by taking the arithmetic mean of each variable across 17 GCMs. (a) Annual mean temperature; (b) Mean diurnal temperature range: mean of monthly (max temperature − min temperature); (c) Isothermality: mean diurnal range/temperature annual range; (d) Temperature seasonality: temperature standard deviation; (e) Max temperature of warmest month; (f) Minimum temperature of coldest month; (g) Temperature annual range: max temperature of warmest month − min temperature of coldest month; (h) Mean temperature of wettest quarter; (i) Mean temperature of driest quarter; (j) Mean temperature of warmest quarter; (k) Mean temperature of coldest quarter; (l) Annual precipitation in mm; (m) Precipitation of wettest month; (n) Precipitation of driest month; (o) Precipitation seasonality (coefficient of variation); (p) Precipitation of wettest quarter; (q) Precipitation of driest quarter; (r) Precipitation of warmest quarter; (s) Precipitation of coldest quarter; (t) Alpine-tundra biome; (u) Cliffs with 2 km buffer. GCM = global climate model; NaN = not a number (area beyond study extent with no data); RCP = representative concentration pathway.

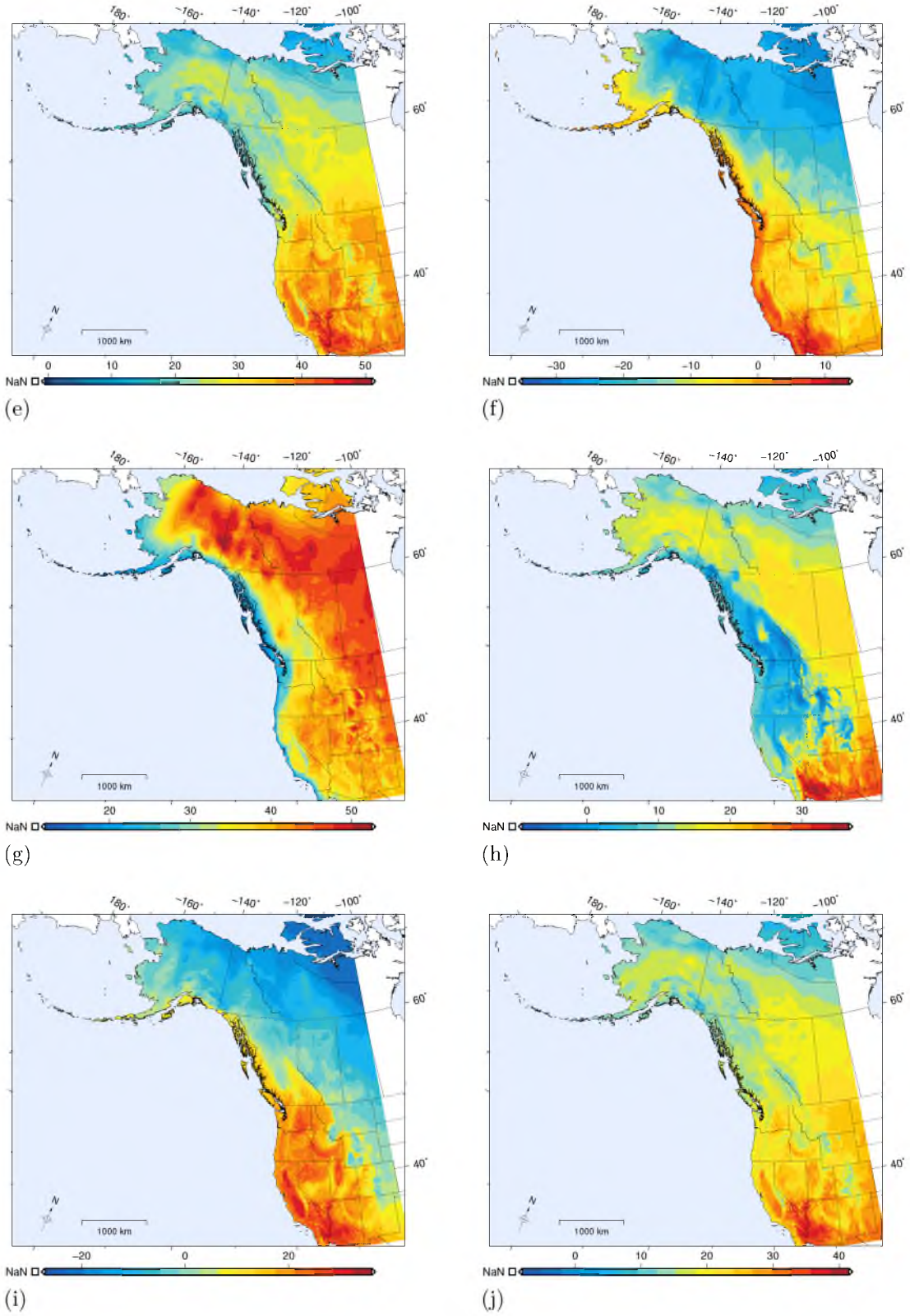


Figure 13. continued.

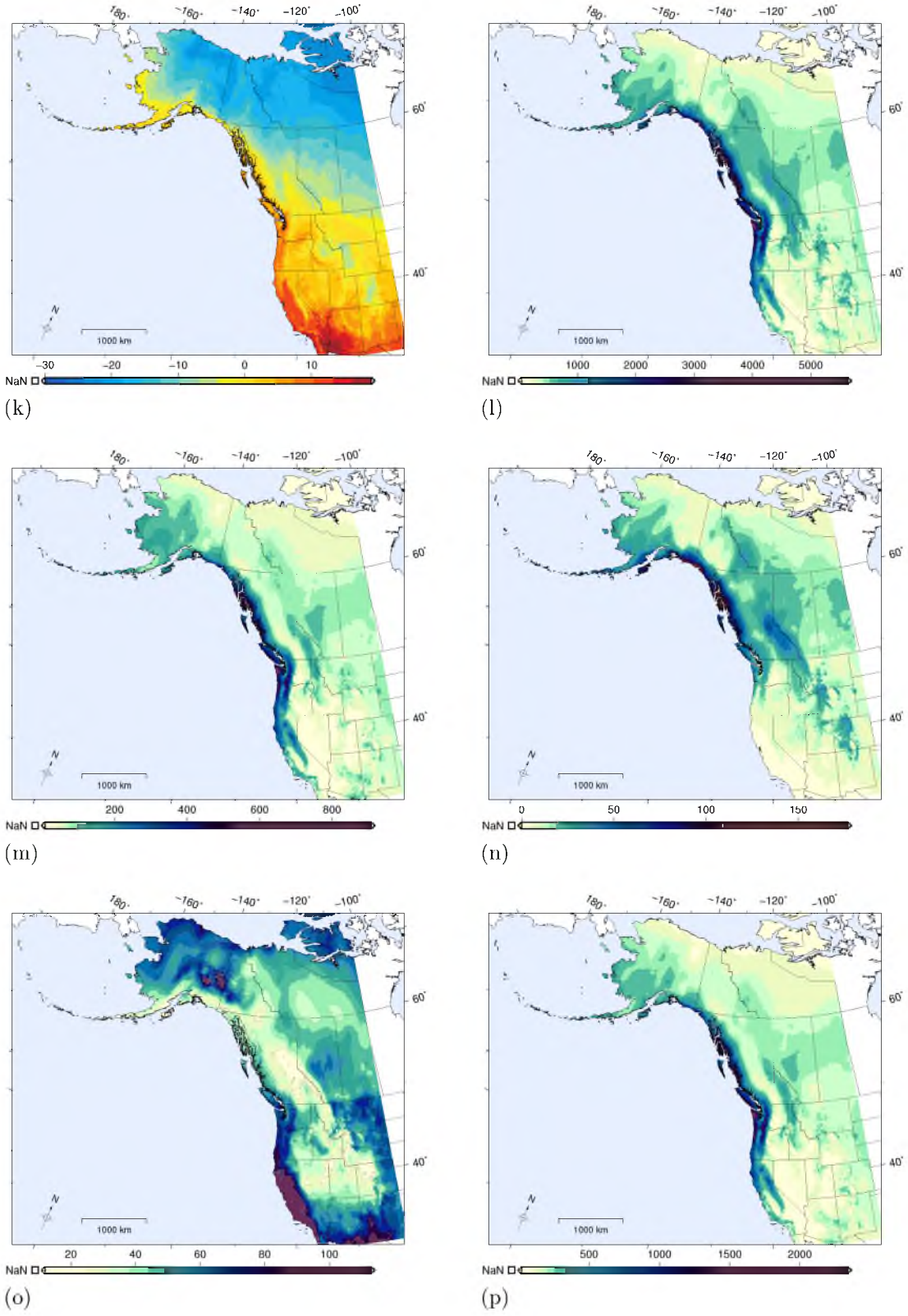
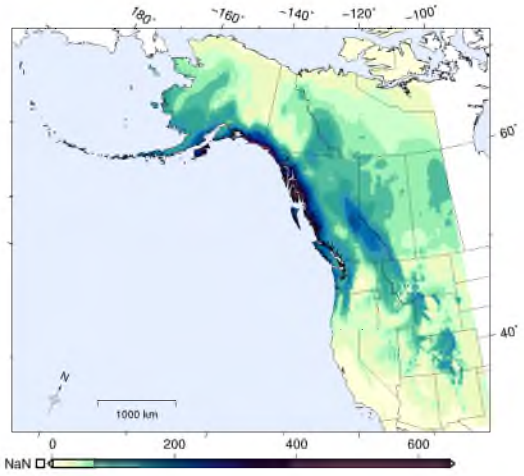
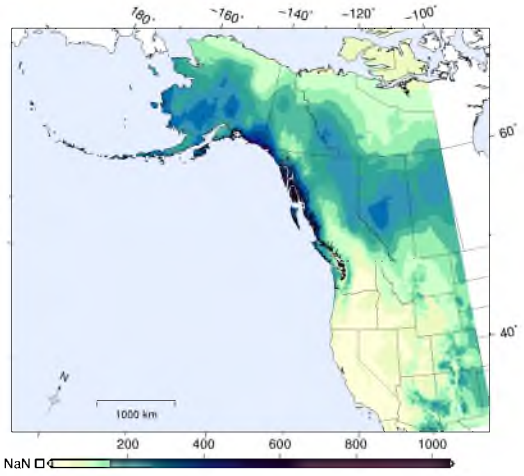


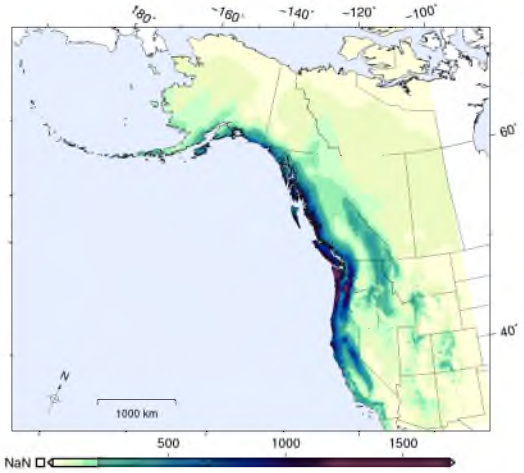
Figure 13. continued.



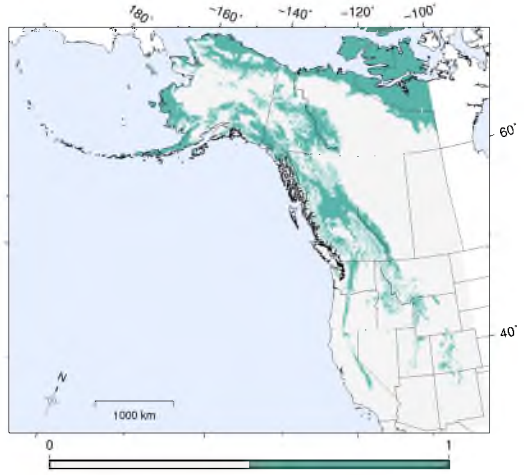
(q)



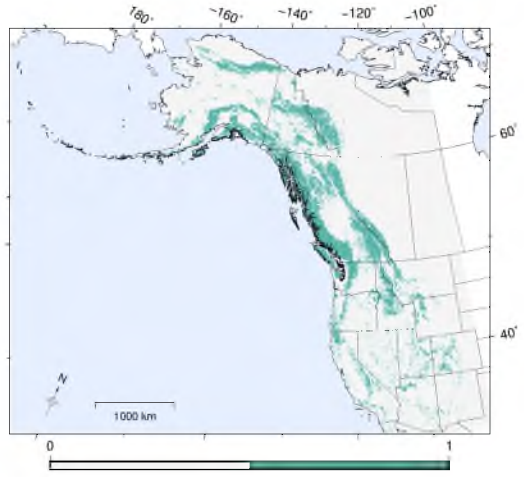
(r)



(s)



(t)



(u)

Figure 13. continued.

REFERENCES

American Ornithologists' Union (1993). Thirty-ninth supplement to the American Ornithologists' Union check-list of North American birds. *The Auk 110*, 675–682.

Araújo, M. B., & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, *16*(6), 743–753.

Beale, C. M., Lennon, J. J., & Gimona, A. (2008). Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences*, *105*(39), 14908–14912.

Beever, E. A., Brussard, P. F., & Berger, J. (2003). Patterns of apparent extirpation among isolated populations of pikas (*Ochotona princeps*) in the Great Basin. *Journal of Mammalogy*, *84*(1), 37–54.

Beever, E. A., Ray, C., Mote, P. W., & Wilkening, J. L. (2010). Testing alternative models of climate-mediated extirpations. *Ecological Applications*, *20*(1), 164–178.

Beever, E. A., Wilkening, J. L., McIvor, D. E., Weber, S. S., & Brussard, P. F. (2008). American pikas (*Ochotona princeps*) in northwestern Nevada: A newly discovered population at a low-elevation site. *Western North American Naturalist*, *68*(1), 8–14.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Brown, D. E. (1998). *A classification of North American biotic communities*. Salt Lake City, UT: University of Utah Press.

Buckley, L. B. (2008). Linking traits to energetics and population dynamics to predict lizard ranges in changing environments. *The American Naturalist*, *171*(1), E1–E19

Cairns, D. M., & Moen, J. (2004). Herbivory influences tree lines. *Journal of Ecology*, *92*(6), 1019–1024.

Chapin III, F. S., Shaver, G. R., Giblin, A. E., Nadelhoffer, K. J., & Laundre, J. A. (1995). Responses of arctic tundra to experimental and observed changes in climate. *Ecology*, *76*(3), 694–711.

Chesser, R. T., Banks, R. C., Barker, F. K., Cicero, C., Dunn, J. L., Kratter, A. W., . . . Winker, K. (2013). Fifty-fourth supplement to the American Ornithologists' Union check-list of North American birds. *The Auk*, *130*(3), 558–572.

Colwell, R. K., & Rangel, T. F. (2009). Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences*, *106*(Suppl. 2), 19651–19658.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., . . . Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*(15), 2031–2064.

Davis, A. J., Jenkinson, L. S., Lawton, J. H., Shorrocks, B., & Wood, S. (1998). Making mistakes when predicting shifts in species range in response to global warming. *Nature*, *391*(6669), 783–786.

Dirnböck, T., Dullinger, S., & Grabherr, G. (2003). A regional impact assessment of climate and land-use change on alpine vegetation. *Journal of Biogeography*, *30*(3), 401–417.

Dirnböck, T., Essl, F., & Rabitsch, W. (2011). Disproportional risk for habitat loss of high-altitude endemic species under climate change. *Global Change Biology*, *17*(2), 990–996.

Drovetski, S. V., Zink, R. M., & Mode, N. A. (2009). Patchy distributions belie morphological and genetic homogeneity in Rosy-Finches. *Molecular Phylogenetics and Evolution*, *50*(3), 437–445.

eBird. (2014). *ebird basic dataset.* Ithaca, New York. [Data file]. Retrieved Version: EBD_relAug–2014, from http://ebird.org/ebird/data/download

Elith, J., & Graham, C. H. (2009). Do they? How do they? WHY do they differ?

On finding reasons for differing performances of species distribution models. *Ecography*, *32*(1), 66–77.

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., . . . Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129–151.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, *17*(1), 43–57.

Elton, C. S. (1927). *Animal ecology*. Chicago, IL: University of Chicago Press.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(01), 38–49.

Fitzpatrick, M. C., & Hargrove, W. W. (2009). The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation*, *18*(8), 2255–2261.

Foreyt, W. J. (1989). Fatal *Pasteurella haemolytica* pneumonia in Bighorn Sheep after direct contact with clinically normal domestic sheep. *American Journal of Veterinary Research*, *50*(3), 341–344.

Foreyt, W. J., & Jessup, D. A. (1982). Fatal pneumonia of Bighorn Sheep following association with domestic sheep. *Journal of Wildlife Diseases*, *18*(2), 163–168.

Foreyt, W. J., Snipes, K. P., & Kasten, R. W. (1994). Fatal pneumonia following inoculation of healthy Bighorn Sheep with *Pasteurella haemolytica* from healthy domestic sheep. *Journal of Wildlife Diseases*, *30*(2), 137–145.

Franklin, J. (2010). *Mapping species distributions: Spatial inference and prediction*. New York, NY: Cambridge University Press.

Freeman, E. A., & Moisen, G. G. (2008a). A comparison of the performance of

threshold criteria for binary classification in terms of predicted prevalence and Kappa. *Ecological Modelling*, *217*(1), 48–58.

Freeman, E. A., & Moisen, G. G. (2008b). PresenceAbsence: An R package for presence absence analysis. *Journal of Statistical Software*, *23*(11), 1–31. Retrieved from http://www.jstatsoft.org/v23/i11

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 1189–1232.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378.

GBIF. (2012). *Recommended practices for citation of the data published through the GBIF Network.* Retrieved Version: 1.0 (Authored by Vishwas Chavan), from http://links.gbif.org/gbif_best_practice_data_citation_en_v1

GDAL Development Team. (2014). GDAL - Geospatial Data Abstraction Library, Version 1.11.1 [Computer software]. Retrieved from http://www.gdal.org

George, J. L., Martin, D. J., Lukacs, P. M., & Miller, M. W. (2008). Epidemic pasteurellosis in a Bighorn Sheep population coinciding with the appearance of a domestic sheep. *Journal of Wildlife Diseases*, *44*(2), 388–403.

GRASS Development Team. (2012). Geographic Resources Analysis Support System (GRASS GIS) Software, Version 6.4.4 [Computer software]. Retrieved from http://grass.osgeo.org

Grinnell, J. (1917). The niche-relationships of the California thrasher. *The Auk*, 427–433.

Gross, J. E., Singer, F. J., & Moses, M. E. (2000). Effects of disease, dispersal, and area on Bighorn Sheep restoration. *Restoration Ecology*, *8*(4S), 25–37.

Guisan, A., Edwards Jr, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling*, *157*(2), 89–100.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2), 147–186.

Heikkinen, R. K., Luoto, M., Virkkala, R., Pearson, R. G., & Körber, J.-H. (2007). Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, *16*(6), 754–763.

Hijmans, R. J. (2012). Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology*, *93*(3), 679–688.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978.

Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2014). dismo: An R package for species distribution modeling, version 1.0-5 [Computer software]. Retrieved from http://CRAN.R-project.org/package=dismo

Hirzel, A., Helfer, V., & Metral, F. (2001). Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, *145*(2), 111–121.

Huntley, B. (1991). How plants respond to climate change: Migration rates, individualism and the consequences for plant communities. *Annals of Botany*, *67*(Suppl. 1), 15–22.

Huntley, B., & Webb III, T. (1989). Migration: Species' response to climatic variations caused by changes in the Earth's orbit. *Journal of Biogeography*, 5–19.

Hutchinson, G. E. (1957). Concluding remarks. In *Cold Spring Harbor Symposia on Quantitative Biology* (Vol. 22, pp. 415–427).

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York, NY: Springer.

Johnson, R. E. (1973). *The biosystematics of the avian genus* Leucosticte *in North America* (PhD thesis, University of California, Berkeley). Unpublished.

Johnson, R. E. (2002). *Black Rosy-Finch (*Leucosticte atrata*)* (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology: The Birds of North America Online. Retrieved from http://bna.birds.cornell.edu/bna/species/678

Johnson, R. E., Hendricks, P., Pattie, D. L., & Hunter, K. B. (2000). *Brown-*

capped Rosy-Finch (Leucosticte australis) (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology: The Birds of North America Online. Retrieved from http://bna.birds.cornell.edu/bna/species/536

Kearney, M. (2006). Habitat, environment and niche: What are we modelling? *Oikos*, *115*(1), 186–191.

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. *Ecology Letters*, *12*(4), 334–350.

Klanderud, K., & Birks, H. J. B. (2003). Recent increases in species richness and shifts in altitudinal distributions of Norwegian mountain plants. *The Holocene*, *13*(1), 1–6.

Körner, C. (1998). A re-assessment of high elevation treeline positions and their explanation. *Oecologia*, *115*(4), 445–459.

Kullman, L. (2002). Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes. *Journal of Ecology*, *90*(1), 68–77.

Lawler, J. J., White, D., Neilson, R. P., & Blaustein, A. R. (2006). Predicting climate-induced range shifts: Model differences and model reliability. *Global Change Biology*, *12*(8), 1568–1584.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22. Retrieved from http://CRAN.R-project.org/doc/Rnews/

Lobo, J. M. (2008). More complex distribution models or more representative data? *Biodiversity Informatics*(5), 14–19.

Lüthi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., . . . Stocker, T. F. (2008). High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature*, *453*(7193), 379–382.

MacArthur, R. A., & Wang, L. C. (1973). Physiology of thermoregulation in the pika, *Ochotona princeps*. *Canadian Journal of Zoology*, *51*(1), 11–16.

Macdougall-Shackleton, S. A., Johnson, R. E., & Hahn, T. P. (2000). *Gray-crowned Rosy-Finch (*Leucosticte tephrocotis*)* (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology: The Birds of North America Online. Retrieved

from http://bna.birds.cornell.edu/bna/species/559

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, *83*(8), 2248–2255.

MacKenzie, D. I., Nichols, J. D., Sutton, N., Kawanishi, K., & Bailey, L. L. (2005). Improving inferences in population studies of rare species that are detected imperfectly. *Ecology*, *86*(5), 1101–1113.

McPherson, J., Jetz, W., & Rogers, D. J. (2004). The effects of species' range sizes on the accuracy of distribution models: Ecological phenomenon or statistical artifact? *Journal of Applied Ecology*, *41*(5), 811–823.

Morin, X., & Thuiller, W. (2009). Comparing niche-and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology*, *90*(5), 1301–1313.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., . . . Wilbanks, T. J. (2010). The next generation of scenarios for climate change research and assessment. *Nature*, *463*(7282), 747–756.

Mountain Research Initiative EDW Working Group. (2015). Elevation-dependent warming in mountain regions of the world. *Nature Climate Change*, *5*(5), 424–430.

Oksanen, L., & Ranta, E. (1992). Plant strategies along mountain vegetation gradients: A test of two theories. *Journal of Vegetation Science*, *3*(2), 175–186.

Olofsson, J., Oksanen, L., Callaghan, T., Hulme, P. E., Oksanen, T., & Suominen, O. (2009). Herbivores inhibit climate-driven shrub expansion on the tundra. *Global Change Biology*, *15*(11), 2681–2693.

Parmesan, C. (1996). Climate and species' range. *Nature*, *382*(6594), 765.

Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J. K., Thomas, C. D., Descimon, H., . . . Warren, M. (1999). Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, *399*(6736), 579–583.

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), 37–42.

Partners in Flight. (2013). *Partners in flight science committee. Population estimates database, Version 2.0.* [Data file]. Retrieved from http://rmbo.org/pifpopestimates

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography*, *12*(5), 361–371.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3), 231–259.

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161–175.

Porfirio, L. L., Harris, R. M., Lefroy, E. C., Hugh, S., Gould, S. F., Lee, G., . . . Mackey, B. (2014). Improving the use of species distribution models in conservation planning and management under climate change. *PloS One*, *9*(11), e113749.

Post, E., & Pedersen, C. (2008). Opposing plant community responses to warming with and without herbivores. *Proceedings of the National Academy of Sciences*, *105*(34), 12353–12358.

Prentice, I. C., Bartlein, P. J., & Webb III, T. (1991). Vegetation and climate change in eastern North America since the Last Glacial Maximum. *Ecology*, 2038–2056.

R Core Team. (2014). R: A language and environment for statistical computing [Computer software]. Vienna, Austria. Retrieved from http://www.R-project.org/

Rehfeldt, G. E., Crookston, N. L., Sáenz-Romero, C., & Campbell, E. M. (2012). North American vegetation model for land-use planning in a changing climate: A solution to large classification problems. *Ecological Applications*, *22*(1), 119–141.

Ridgeway, G. (2013). gbm: Generalized boosted regression models [Computer

software]. Retrieved from http://CRAN.R-project.org/package=gbm (R package version 2.1)

Root, T. L., Price, J. T., Hall, K. R., Schneider, S. H., Rosenzweig, C., & Pounds, J. A. (2003). Fingerprints of global warming on wild animals and plants. *Nature*, *421*(6918), 57–60.

Rota, C. T., Fletcher, R. J., Evans, J. M., & Hutto, R. L. (2011). Does accounting for imperfect detection improve species distribution models? *Ecography*, *34*(4), 659–670.

Sekercioglu, C. H., Schneider, S. H., Fay, J. P., & Loarie, S. R. (2008). Climate change, elevational range shifts, and bird extinctions. *Conservation Biology*, *22*(1), 140–150.

Sibley, C. G., & Monroe, B. L. (1990). *Distribution and taxonomy of birds of the world*. New Haven, CT: Yale University Press.

Simpson, W. G. (2009). American pikas inhabit low-elevation sites outside the species' previously described bioclimatic envelope. *Western North American Naturalist*, *69*(2), 243–250.

Singer, F. J., Zeigenfuss, L. C., & Spicer, L. (2001). Role of patch size, disease, and movement in rapid extinction of Bighorn Sheep. *Conservation Biology*, *15*(5), 1347–1354.

Smith, A. T. (1974). The distribution and dispersal of pikas: Influences of behavior and climate. *Ecology*, 1368–1376.

Speed, J. D., Austrheim, G., Hester, A. J., & Mysterud, A. (2010). Experimental evidence for herbivore limitation of the treeline. *Ecology*, *91*(11), 3414–3420.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., . . . Midgley, P. M. (2013). Climate change 2013: The physical science basis. *Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5)*.

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, *142*(10), 2282–2292.

Sutter, J. V., Andersen, M. E., Bunnell, K. D., Canning, M. F., Clark, A. G., Dolsen, D. E., & Howe, F. P. (2005). Utah comprehensive wildlife conservation strategy (CWCS). *Utah Division of Wildlife Resources, Publication*(05-19).

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498.

Thuiller, W. (2007). Biodiversity: Climate change and the ecologist. *Nature*, *448*(7153), 550–552.

Thuiller, W., Brotons, L., Araújo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, *27*(2), 165–172.

United States Geological Survey. (2010). *USGS Global Multi-resolution Terrain Elevation Data (GMTED) product.* [Data file]. Retrieved from http://topotools.cr.usgs.gov/gmted_viewer/

Virtanen, R., Luoto, M., Rämä, T., Mikkola, K., Hjort, J., Grytnes, J.-A., & Birks, H. J. B. (2010). Recent vegetation changes at the high-latitude tree line ecotone are controlled by geomorphological disturbance, productivity and diversity. *Global Ecology and Biogeography*, *19*(6), 810–821.

Walker, M. D., Wahren, C. H., Hollister, R. D., Henry, G. H., Ahlquist, L. E., Alatalo, J. M., . . . Wookey, P. A. (2006). Plant community responses to experimental warming across the tundra biome. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(5), 1342–1346.

Webb III, T., & Bartlein, P. J. (1992). Global changes during the last 3 million years: Climatic controls and biotic responses. *Annual Review of Ecology and Systematics*, 141–173.

Wessel, P., Smith, W. H., Scharroo, R., Luis, J., & Wobbe, F. (2013). Generic mapping tools: Improved version released. *Eos, Transactions American Geophysical Union*, *94*(45), 409–410.

Williams, J. W., & Jackson, S. T. (2007). Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, *5*(9),

475–482.

Williams, J. W., Jackson, S. T., & Kutzbach, J. E. (2007). Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences*, *104*(14), 5738–5742.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(1), 3–36.

Zink, R. M., Rohwer, S., Andreev, A. V., & Dittmann, D. L. (1995). Trans-Beringia comparisons of mitochondrial DNA differentiation in birds. *Condor*, 639–649.